Integration of quantitative and molecular genetic approaches to improve characteristics associated with pig welfare

Dagmar Nicoline Reinhildis Gertrud Kapell

Thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh

2010

Declaration

I declare that this thesis is my own composition and that the research described in it is my own work, except where acknowledged. Neither this work, nor any part of it, has previously been submitted for any other degree or professional qualification.

D.N.R.G. Kapell
2010

Publications

Refereed papers

KAPELL, DNRG, ASHWORTH, CJ, WALLING, GA, LAWRENCE, AB, EDWARDS, SA and ROEHE, R, 2009. Estimation of genetic associations between reproduction and production traits based on a sire and dam line with common ancestry. Animal 3: 1354-1362 [Based on Chapter 3]

KAPELL, DNRG, ASHWORTH, CJ, KNAP, PW and ROEHE, R, 2011. Genetic correlations between piglet survival, litter size and birth weight or its variation within litter in sire and dam lines using Bayesian analysis. Livestock Science 135: 215-224 [Based on Chapter 2]

Conference papers

KAPELL, DNRG, ASHWORTH, CJ, WALLING, GA, LAWRENCE, AB, EDWARDS, SA and ROEHE, R, 2008. Estimation of genetic associations between piglet survival and performance traits for sire and dam lines with common ancestry. Communication No. 103 in *Proceedings of the British Society of Animal Science 2008, Scarborough, United Kingdom*. [Based on Chapter 3]

KAPELL, DNRG, ASHWORTH, CJ, KNAP, PW and ROEHE, R, 2009. Genetic correlations between piglet birth weight, its variation and survival in different dam and sire lines using Bayesian analysis. Communication No. 59 in *Proceedings of the British Society of Animal Science 2009, Southport, United Kingdom.* [Based on Chapter 2]

KAPELL, DNRG, SORENSEN, D, SU, G, JANSS, LLG, ASHWORTH, CJ and ROEHE, R, 2009. Performance of genomic selection for traits in mice using Bayesian multi-marker models. Communication No. 28-04 *European Federation of Animal Science, Barcelona, Spain.* [Based on Chapter 4]

Abstract

The aims of this thesis were to investigate whether characteristics associated with animal welfare are genetically and genomically determined by using quantitative and molecular genetic approaches and to develop strategies indicating how these traits could be used in breeding programmes. Two traits that are closely related to animal welfare and associated with high socio-economic values are piglet survival at birth and aggressive behaviour in pigs. Piglet survival traits were analysed based on quantitative Bayesian approaches using phenotypic and pedigree information only, while aggressive behaviour was analysed based on molecular genetic approaches such as genome-wide association studies and genomic selection using additionally a dense panel of single nucleotide polymorphisms (SNP). The latter approach was validated using behavioural traits related to welfare characteristics in a well-documented mouse data set.

Selection for piglet survival at birth is expected to be effective, because all lines and breeds in this thesis showed considerable variation for this trait and relatively high heritabilities, particularly in lines with low average birth weight. Maternal heritabilities of individual birth weight were mostly at moderate magnitude and thus of great interest for selection. The genetic correlations between piglet survival and birth weight indicated that selection for either individual or average birth weight or variation of birth weight within litter would indirectly increase survival. The genetic associations of piglet survival with economically important (re)production traits are of great importance for breeding organisations. Undesirable genetic correlations between piglet survival and (re)production traits were generally of low magnitude, so that simultaneous improvement of all traits could be achieved. A comparison of five breeds and lines showed that differences in genetic parameters between breeds and lines can be substantial and no single selection strategy would be optimal for all. A unique study of a sire and a dam line originating from one breed but selected for more than 25 years with different breeding goals demonstrated how selection pressure can alter the genetic parameters over years. Breeding organisations should

therefore consider selection strategies per breed or line individually to achieve maximum overall improvement.

This study gives new insight into the use of genomic selection for traits associated with animal welfare. It is one of the first to present estimates for linkage disequilibrium in the pig using a new 60K SNP panel and the first to evaluate the efficiency of genomic selection against aggressive behaviour in pigs. Genomic selection showed a high predictive ability in comparison to traditional polygenic selection. It was especially advantageous for traits with lower heritabilities. In particular in situations where little family information was available, the performance of polygenic selection was low and genomic selection increased the performance considerably. Reducing the number of SNPs did not significantly change the performance of genomic selection. The consistently high performance across models indicates that low-density SNP panels may be sufficient to ensure a high efficiency of genomic selection, thus reducing the high costs associated with genotyping in pig breeding with its short generation interval. To summarize, this thesis has shown how to optimise quantitative and genomic approaches to improve animal welfare related characteristics efficiently in pig breeding programmes.

Acknowledgements

First and foremost I would like to thank my main supervisor, Rainer Roehe, and co-supervisor, Cheryl Ashworth, for their ideas and support and for sharing their knowledge with me. I am especially grateful to Rainer for giving me the chance to do this PhD and for his guidance throughout these four years.

Sponsorship for this PhD was provided by the SABRETRAIN Project, funded by the Marie Curie Host Fellowships for Early Stage Research Training as part of the 6th Framework Programme of the European Commission.

Next I would like to thank Grant Walling from JSR Genetics Ltd and Pieter Knap from PIC International for providing the data sets for my research on piglet survival and for their input and ideas for the respective parts of my work. Similarly, I would like to thank Alistair Lawrence, Simon Turner and Rick D'Eath from SAC for providing the data sets and their input for my research on aggressive behaviour in pigs.

I would also like to thank Daniel Sorensen, Guosheng Su and Luc Janss from the Faculty of Agricultural Sciences (Aarhus University) for giving me the opportunity to conduct part of my research on genomic selection in mice with them in Viborg. I gratefully acknowledge the Wellcome Trust Centre for Human Genetics for providing the mouse data. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk).

Acknowledgements

A big thanks goes to the two people I shared most of my office time with, Jenny and Stephanie. Thank you both for telling me repeatedly that my spelling is pants, my grammar sucks and my general knowledge of English is shameful… I could not have done it without you. I would also like to thank my tea-drinking buddies at SAC, for many entertaining discussions and making my time there so much fun.

Met dank aan iedereen in bierenfriet voor het luisteren naar mijn geneuzel al die jaren en om, ondanks boeingsfactor negatief, toch te doen alsof het ze boeide.

Miene dank geit ouch oet noa Monique, ummer doa óm mit te kalle en óm noa mich te loestere, auch al waars te aaf en toe versjtop online. ‘t is altiejd goud óm eemes te höbbe dae ’t zelfde mitmaak, en doaróm ouch precies sjnap wasse mèns.

Miene grootste dank geit oet noa mien ouwesj, die mich alle joare lank óngersjteund höbbe en ummer doa waasse veur mich ès ich ze neudig houw. Zónger uch weur ich neit zo wiejd gekomme en haej ich dit neit kènne doon. Geer höb in mich gegluif, zelfs ès ich zelf het geluif verloare houw. Ich zal uch veur ummer dankbaar zeen.

Abbreviations

95%-HPD	95%-highest posterior density interval
BLUP	best linear unbiased prediction
GS	genomic selection
HWE	Hardy-Weinberg equilibrium
LD	linkage disequilibrium
MAF	minor allele frequency
MAS	marker assisted selection
MSD	mean square difference
PA	predictive ability
PPOR	change in odds from prior to posterior probability
QTL	quantitative trait locus
REML	restricted maximum likelihood
SNP	single nucleotide polymorphism

Traits

ADG	average daily gain
ALBW	average birth weight within litter
BF	backfat thickness
DNRA	delivery of non-reciprocal aggression
DW	percentage of piglets dead from birth till weaning
FB	fecal boli after cue
HC	hematocrit percentage
I75	insulin level at 75 minutes after injection of glucose
IBW	individual piglet birth weight
LSA1	lesion score at mixing anterior region
LSC1	lesion score at mixing central region
LSE1	lesion score at mixing caudal region
LSA2	lesion score post mixing anterior region

LSC2	lesion score post mixing central lesion
LSE2	lesion score post mixing caudal region
MD	muscle depth
NBA	number of piglets born alive
NBT	number of piglets born in total
NW	number of piglets weaned
RA	reciprocal aggression
RNRA	receipt of non-reciprocal aggression
SB	percentage of piglets stillborn
STD	standard deviation of birth weight within litter
SVBL	survival at birth at litter level (%)
SVBP	survival at birth at piglet level (%)
TA	total activity in open field test
TF	time freezing during cue
W10	weight at week 10
W6	weight at week 6
W6m	weight at week 6 (missing)

Table of contents

List of tables

List of figures

Chapter 1 - General Introduction

1.1 Genetics and animal welfare

Animal welfare is of importance to producers (STOTT *et al.* 2005; MORGAN-DAVIES *et al.* 2006) as well as consumers (CRONEY and MILLMAN 2007; TOMA *et al.* 2010). Improvement of animal welfare in our current production systems is of great value for many reasons, including not only economic benefits for the producers, but also moral and ethical concerns, compliance with (inter)national legislation (BORGEN and SKARSTAD 2007; LAWRENCE and STOTT 2009), trade benefits and environmental impact (TOMA *et al.* 2008). Traits that are associated with aspects of animal welfare will have a substantial role in future breeding programmes (e.g. MUIR and CRAIG 1998; KANIS *et al.* 2004; D'EATH *et al.* 2010; RODENBURG *et al.* 2010).

The overall aim of this thesis was to investigate whether characteristics associated with animal welfare are genetically and genomically determined and to develop strategies which indicate how these traits could be used in breeding programmes. This was done by using quantitative and molecular genetic approaches. In this thesis, quantitative genetic approaches were used in the narrow sense as methods using phenotypic performance and all pedigree information by including the additive genetic relationship matrix into the genetic model to predict genetic and environmental parameters of the traits of interest. Molecular genetic approaches were used in the broader sense as methods using genomic information based on high-density single nucleotide polymorphism (SNP) panels, beside phenotypic performance and genetic relationships. These definitions have been used to clearly describe the different methods used in this thesis, although it is recognised that some methods described as molecular genetic approaches can also be described as quantitative approaches.

Two traits that are closely related to animal welfare and associated with high socio-economic values are piglet survival at birth (EDWARDS 2002; KANIS *et al.* 2004; ALONSO-SPILSBURY *et al.* 2007) and aggressive behaviour in pigs (KANIS *et al.* 2004; KANIS *et al.* 2005; D'EATH *et al.* 2010). In this thesis, piglet survival traits are

analysed based on quantitative approaches using Bayesian methodology. In contrast, aggressive behaviour in pigs is analysed based on molecular genetic approaches using genome-wide association studies and genomic selection. This approach has been validated using behavioural traits related to welfare characteristics in a well-documented mouse data set.

1.2 Quantitative genetics

1.2.1 Piglet survival

Sow productivity traits are of major economic importance in the global pig breeding industry, as is shown in studies from Finland (SERENIUS and MUHONEN 2007) and Canada (QUINTON *et al.* 2006). Within sow productivity traits, the relative economic value of piglet survival, expressed as percent of the economic value for litter size can be as high as 14.5% (QUINTON *et al.* 2006). Moreover, the economic value of piglet mortality using a bio-economic model, expressed in Euro per piglet, is € -4.42 compared to € 3.42 per piglet born (SERENIUS and MUHONEN 2007). The number of piglets sold per sow per year plays an important role in the economic viability of a weaner producer. Many breeding goals include selection for a litter size trait, such as number born in total or number born alive. However, studies have shown that this can have a negative impact on piglet survival. For example, piglets of Landrace and Yorkshire breeds, with a total number of piglets born of 14.3 and 13.1 piglets, respectively, have survival rates at birth of 82.6% and 88.7% (SU *et al.* 2007). In the UK, the number of piglets born alive and reared per sow per litter has grown steadily over the years (Figure 1.1).

Internationally, piglet mortality from birth to weaning is in the range of 10% to 25% (reviewed by ALONSO-SPILSBURY *et al.* 2007). In the UK, the mean stillbirth rate is reported to be 6.4% with a mean mortality of live born pigs before weaning of 12.5% (BPEX 2010). Between 2004 and 2006, piglet mortality of live born piglets before weaning in the breeding herd increased slightly. However, this declined in later

years, whereas piglet mortality in the feeding and rearing herd decreased between 2004 and 2009 (Figure 1.2).

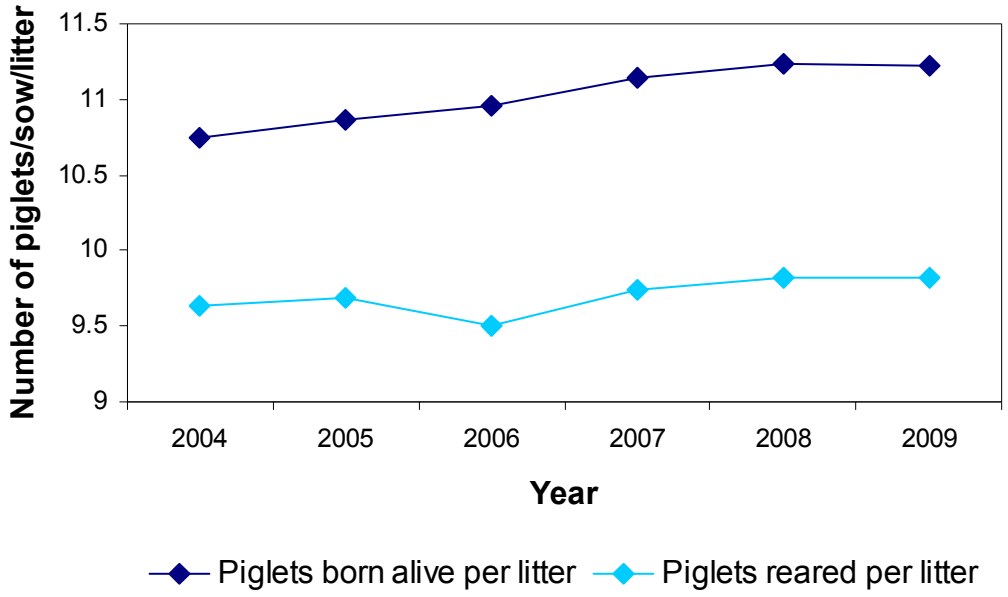


Figure 1.1: Number of piglets born alive and reared. Average number of piglets born alive and reared per sow per litter from 2004 to 2009 in the UK (BPEX 2008; BPEX 2010).

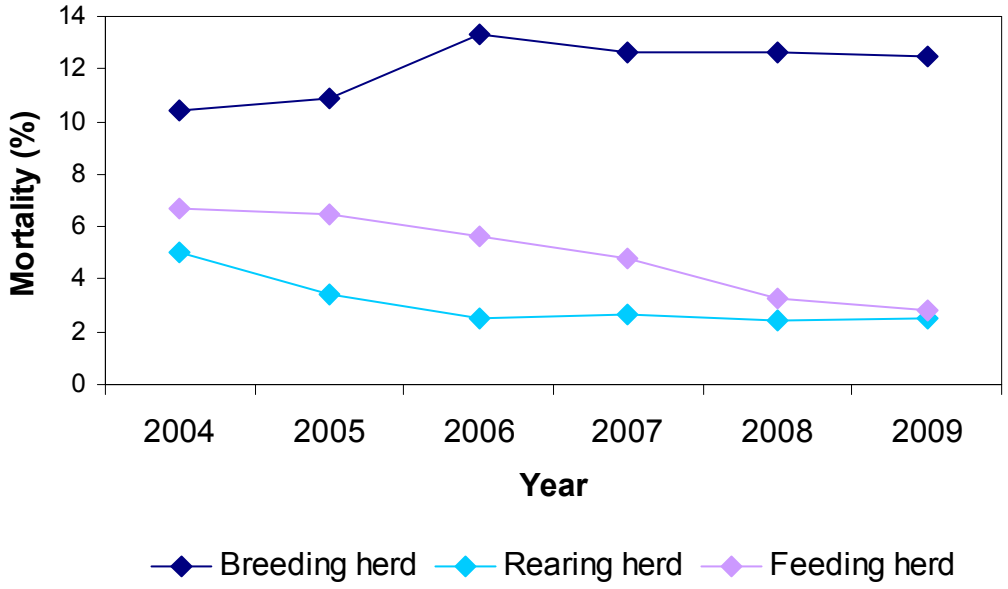


Figure 1.2: Mortality from birth until slaughter. Average mortality of pigs in the breeding, rearing and feeding herd from 2004 to 2009 in the UK (BPEX 2008; BPEX 2010).

This scale of piglet mortality forms a considerable economic loss for pig producers and raises substantial animal welfare concerns. Various factors that contribute to piglet mortality from birth till weaning have been described in the literature, with complex interactions between sow, piglet and environmental factors (Figure 1.3). In addition, BAXTER *et al*. (2008) found novel indicators for postnatal survival, including birth weight, vigour independent of birth weight and the latency to first suckle.

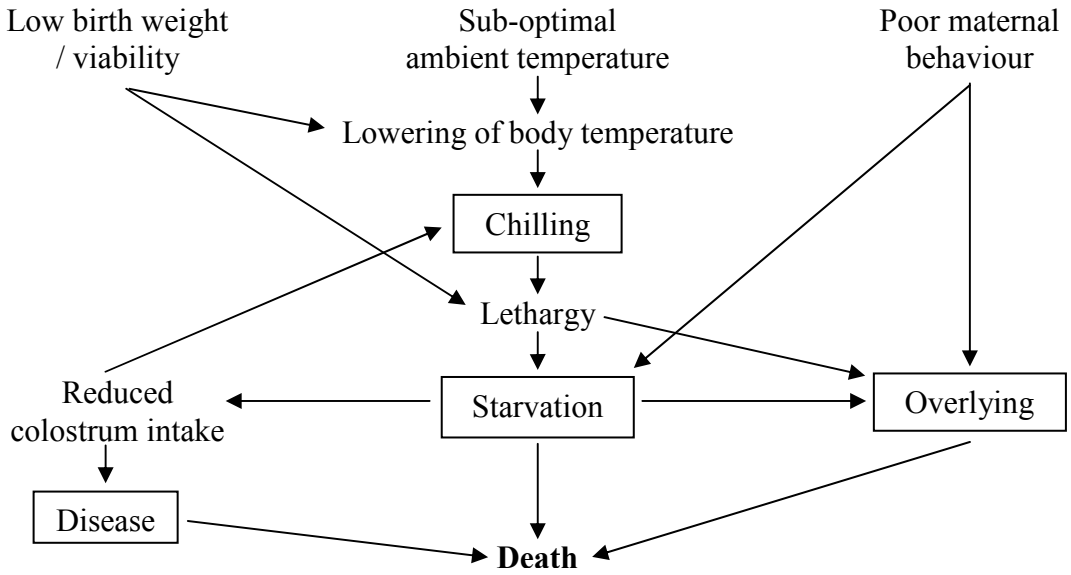


Figure 1.3: Factors contributing to piglet mortality. The complex interactions occurring in the chilling-starvation-overlying-disease complex (from EDWARDS 2002).

Even before birth, the maternal environment influences the development and survival of the oocyte and embryo, for example through the nutritional status of the dam, and changes in nutrient supply can influence the viability of the resulting offspring (reviewed by ASHWORTH *et al.* 2009). During and after birth, maternal behaviour of the sow, which leads to starvation and overlying (crushing), plays an important role in piglet mortality. Crushing is the largest contributor to piglet mortality in outdoor breeding herds and the second largest, after stillbirth, in indoor breeding herds, followed by starvation as third largest in both environments (EDWARDS 2002).

However, heritabilities of maternal behaviour traits, if considered at all, were low, e.g. a heritability of 0.06 for the piglet scream test (reviewed by GRANDINSON 2005). BAXTER *et al.* (2008) found piglet shape and size, body mass index and farrowing birth order to be indicators of piglet mortality at birth. Farrowing traits, such as farrowing duration, birth interval, heterogeneity of birth intervals within litter and degree of birth assistance, show low heritabilities (ranging from 0.01 to 0.12) but moderate to high genetic correlations (ranging from 0.20 to 0.90) with piglet survival at birth (CANARIO *et al.* 2006b). Various characteristics related to birth weight – such as a high variation of birth weight within a litter (MILLIGAN *et al.* 2002) or an increased difference between piglet birth weight and mean litter birth weight (CANARIO *et al.* 2006a) – are associated with an increase in piglet mortality.

Another factor of importance for piglet survival is body condition of the sow, which can be analysed as weight and backfat thickness at farrowing and weight loss and backfat loss during lactation (GRANDINSON *et al.* 2005). Furthermore, male piglets are 1.8 times more likely to be stillborn (CANARIO *et al.* 2006a) and 1.5 times more likely to die pre-weaning (ROEHE and KALM 2000) than female piglets. Efforts to increase piglet survival through changes to husbandry and housing – for example through farrowing crates – have reached their limit in terms of improved survival and have raised public concerns about the welfare implications of confining sows (EDWARDS 2002). Therefore, research described in this thesis emphasised the enhancement of piglet survival through genetic improvement.

1.2.2 Genetic determination of piglet survival

Several studies have investigated the survivability of piglets at birth and from birth till weaning, with emphasis on the estimation of the heritabilities of a range of traits related to piglet survival, for example farrowing survival, pre-weaning survival or total survival (KNOL *et al.* 2002a) or survival at birth, survival during early pre-weaning or survival during late pre-weaning (SU *et al.* 2008). Most studies considered piglet survival at litter level and found heritabilities ranging from 0.02 to 0.11 (GRANDINSON *et al.* 2002; SERENIUS *et al.* 2004b; ARANGO *et al.* 2005;

ROSENDO *et al.* 2007b). Few investigated piglet survival at individual piglet level, with estimates of heritabilities for the direct (maternal) genetic component ranging from 0.02 to 0.11 (0.04 to 0.13) (VAN ARENDONK *et al.* 1996; IBÁÑEZ-ESCRICHE *et al.* 2009b), whereby direct and maternal genetic effects may be difficult to disentangle (WILLHAM 1980). Although the heritabilities are generally low, the genetic variation is large enough to provide improvement through breeding (KNOL *et al.* 2002a; SU *et al.* 2008; ROEHE *et al.* 2009), despite mostly negative estimates of correlations between the direct and maternal genetic component of individual piglet survival, which ranged from -0.56 to 0.15 (VAN ARENDONK *et al.* 1996; IBÁÑEZ-ESCRICHE *et al.* 2009b). The study described in **Chapter 2** analysed piglet survival at birth at litter level as well as piglet level, to obtain further insight into the heritabilities and direct-maternal genetic correlations.

1.2.3 Bayesian analysis

Most studies have examined the survivability of piglets at birth and from birth to weaning using linear models based on a restricted maximum likelihood approach (REML) (KNOL *et al.* 2002a; SERENIUS *et al.* 2004b). However, for binary traits, such as survival (alive or dead), threshold models have been shown to be more appropriate (SORENSEN *et al.* 1995). Software has recently been developed using Bayesian methodology to estimate genetic (co)variances amongst categorical threshold traits, as well as between threshold traits and continuous linear traits. This allows the threshold characteristic of individual piglet survival at birth to be taken into account, and provides an opportunity to examine the precision of estimation of genetic parameters by using Bayesian confidence intervals. Both studies described in **Chapters 2 and 3** use a Bayesian approach, which is an especially powerful statistical tool for the estimation of genetic parameters of threshold traits such as individual piglet survival.

1.2.4 Genetic association of piglet survival with birth weight

Piglet survival is phenotypically correlated with piglet birth weight – whereby a higher birth weight is associated with a higher postnatal survival – and may be

indirectly improved by using the latter trait (ROEHE 1999; ROEHE and KALM 2000). Moreover, little is known about the genetic associations between direct or maternal effects of piglet survival and piglet birth weight or its variation within litter. Nonetheless, these may have substantial potential to be exploited for genetic improvement of piglet survival. A higher within-litter birth weight variation is associated with lower survival, independent of the litter size, while the mean weaning weight is positively correlated with mean piglet birth weight of the litter, and negatively correlated with litter size (MILLIGAN *et al.* 2002). Birth weight and survival at birth have often only been analysed at the litter level, as a trait of the dam. Few studies took the genetic component of the individual piglet into account (GRANDINSON *et al.* 2002; BOUQUET *et al.* 2006; SU *et al.* 2008; ROEHE *et al.* 2009). However, individual piglet birth weight may be closely related to piglet survival and might therefore be used to indirectly improve piglet survival at birth. In **Chapter 2,** the genetic background of survival, litter size and weight traits and their correlations at litter level as well as at individual piglet level is examined and discussed.

1.2.5 Genetic association of piglet survival with reproduction traits

Lower litter sizes at weaning involve a considerable economic loss for pig producers. In 2009, the gross margin per pig for weaner producers was £ 28.12 at 19.7 pigs sold per sow per year (BPEX 2010). Selection for piglet survival is even more important because unfavourable genetic correlations have been estimated between total number of piglets born and neonatal survival at -0.46 to -0.06 (SERENIUS *et al.* 2004b; BOUQUET *et al.* 2006; ROSENDO *et al.* 2007b; SU *et al.* 2007) or between total number of piglets born and postnatal survival at -0.52 to -0.07 (SERENIUS *et al.* 2004b; SU *et al.* 2007). The current selection pressure on increased litter size is likely to increase the mortality rate, if not accompanied by selection for piglet survival, as for example, reported by Danish and a Dutch pig breeding organisations (KNAP 2008). Therefore, selection strategies have to be developed to improve both litter size and survival simultaneously.

Correlations between survival traits and reproduction traits show contradictory results in the literature (SERENIUS *et al.* 2004b; ROSENDO *et al.* 2007b; SU *et al.* 2007). Traits present in most studies were birth weight and number born alive or number born in total. Generally, these studies investigate maternal genetic effects only, or combine direct and maternal genetic effects (SU *et al.* 2008). Only a few studies include other effects such as nurse sow effects (KNOL *et al.* 2002a) or service sire effects (CHEN *et al.* 2003). In **Chapter 2,** genetic associations between piglet survival at birth and number born in total at litter level are investigated, in addition to the effect of adjustment of individual piglet survival for number born in total on its genetic estimates. Furthermore, in **Chapter 3,** genetic associations of piglet survival at birth with number of piglets born alive and number of piglets born in total are estimated at litter level.

1.2.6 Dam lines versus sire lines

Pig breeding aims to improve total genetic merit by including all traits of economic value into the breeding goal. Pig breeding programmes distinguish between sire and dam lines with different breeding goals for several important reasons. Firstly, a differentiation in breeding goals allows emphasising selection pressure in dam lines on reproduction performance and in sire lines on production performance. Growing-finishing animals are the products of three- or four-way crosses between animals from sire and dam lines, thus allowing the exploitation of heterosis. In addition, the separation of the lines helps to reduce the potential effect of negative genetic correlations between reproduction and production traits on selection response by emphasising selection pressure mostly on one of these traits.

Dam lines produce highly productive mothers of the slaughter pigs. They are bred with emphasis on reproductive traits such as litter size, piglet survival, maternal behaviour, rebreeding interval or longevity. From 2005 to 2009, not only the mean number of piglets born alive and reared increased per sow per litter (Figure 1.1). The number of litters per sow per year increased from 2.22 to 2.25 and the number of piglets reared per sow per year increased from 21.5 to 22.2 (BPEX 2010).

Sire lines produce the sires of the slaughter animals. The emphasis is on production traits such as growth rate, leanness, feed efficiency and meat quality. Between 2005 and 2009, the average daily gain in feeding herds increased from 639 g to 819 g and the lean meat percentage increased from 61.1% to 61.3% (BPEX 2010). A high emphasis on production traits may compromise fitness traits and thus have a negative impact on reproductive abilities (RAUW *et al.* 1998; KNAP and RAUW 2008).

Research has shown that breeding goal differences, with emphasis on reproduction traits in dam lines and on production traits in sire lines, may result in different genetic parameters between lines (KNOL *et al.* 2002a). This structure of breeding programmes is of major importance for the interpretation of the results from **Chapters 2 and 3**. Specifically, **Chapter 2** describes research on the genetic background of traits in five different lines, two sire lines and three dam lines, which originated from different breeds. In contrast, **Chapter 3** presents the research on the difference in the genetic background of traits between a sire line and a dam line originating from the same breed. This means that the change of parameters due to 25 years of selection with different breeding goals in a sire and dam line that originated from the same breed is examined. To the best of my knowledge, at present no studies are available comparing how parameters in sire and dam lines changed over years due to different selection strategies.

1.2.7 Genetic associations of piglet survival with production traits

There are only a few studies available in the literature estimating the correlations between piglet survival traits and production traits. Two production traits that are commonly considered are backfat thickness and growth rate. Pre-weaning survival shows moderately positive correlations with average daily feed intake, backfat thickness and lipid deposition, but a strong negative correlation with residual feed intake (KNOL 2001). Similarly, low to moderately favourable correlations have been estimated between number of piglets born dead and average daily gain (SERENIUS *et al.* 2004a) and days to reach 113.5 kg (ARANGO *et al.* 2005). However, low to moderately unfavourable genetic correlations have been found between number of

piglets born dead and carcass traits such as lean and fat percentage (SERENIUS *et al.* 2004a) and ultrasound backfat thickness (ARANGO *et al.* 2005). Estimates for genetic correlations between number born alive and growth rate traits or backfat thickness are generally low and contradictory between studies (HERMESCH *et al.* 2000c; ARANGO *et al.* 2005). Genetic correlations of number born in total with performance traits – such as average daily gain and feed to gain ratio – are generally low and, depending on breed, favourable or unfavourable (SERENIUS *et al.* 2004a). In **Chapter 3**, the influence of selection on the genetic correlations of piglet survival at birth with average daily gain, backfat thickness and muscle depth is studied for a sire and dam line originating from a common breed.

1.3 Molecular genetics

1.3.1 Animal Aggression

Aggressive behaviour plays an important role in many situations such as defence against predators (QUINN and UETA 2008), acquisition of food (DRUMMOND 2001; ASHLEY 2007), establishment of a social hierarchy (LANGBEIN and PUPPE 2004; ASHLEY 2007) or mating (SMALL 1988). It was a key factor in domestication of animals, with aggressive animals less likely to be used for breeding purposes (PRICE 1984). Aggressive behaviour in animals has been observed in many species. Model species, including Drosophila (EDWARDS *et al.* 2006), mice (BRODKIN *et al.* 2002) and rats (ALBERT *et al.* 2009), have been studied extensively and have been selected successfully for increased or decreased aggression for many generations. Dogs have been selected for centuries for different temperaments, and the genetic background of aggressive behaviour, towards humans as well as towards conspecifics, has been studied widely (PÉREZ-GUISADO *et al.* 2006; LIINAMO *et al.* 2007; VAN DEN BERG *et al.* 2008; TAKEUCHI *et al.* 2009). In livestock species, evidence for genetic components has been found for aggressive behaviour at handling – for example in cattle (PHOCAS *et al.* 2006) or pigs (GRANDINSON *et al.* 2003) – or aggressive behaviour towards conspecifics – for example in chickens (BUITENHUIS *et al.* 2003a;

BUITENHUIS *et al.* 2003b) or pigs (BREUER *et al.* 2005) – or aggressive maternal behaviour – for example in pigs (KNAP and MERKS 1987; CHEN *et al.* 2009).

1.3.2 Aggressive behaviour at mixing in pigs

In commercial pig production, unrelated animals are repeatedly mixed at different stages in their life, e.g. sows at different stages in the reproductive cycle, piglets at weaning or all animals for transport to the slaughterhouse. Establishing a new social hierarchy at mixing takes time and effort, and is associated with increased aggression (AREY 1999). Aggression involves stress and the skin lesions open routes for infection, thus compromising the welfare of the animal. Efforts to reduce aggression at mixing – by, for example, greater space allowance, barriers, provision of ad libitum feed, straw bedding or chemical intervention – showed limited or no improvement (reviewed by AREY and EDWARDS 1998). Recently, research has been carried out to improve aggressive behaviour at mixing through genetic improvement.

Aggressive behaviour of sows at mixing displays low to moderate heritabilities, at 0.04 to 0.24 (LØVENDAHL *et al.* 2005), with slightly higher heritabilities for aggression, at 0.08 to 0.48, and lesion score traits, at 0.11 to 0.43, in slaughter pigs (TURNER *et al.* 2006a; TURNER *et al.* 2008; D'EATH *et al.* 2009; TURNER *et al.* 2009) and the genetic variation is large enough to achieve improvement through breeding. Aggressive behaviour of sows is not significantly correlated with maternal behaviour (LØVENDAHL *et al.* 2005), nor are lesion scores significantly correlated with growth rate or backfat thickness (TURNER *et al.* 2006a). Behavioural data are often costly and difficult to measure, resulting in low numbers of observations and large standard errors of estimates for genetic parameters for the aforementioned studies.

1.3.3 Genome-wide association studies

Genome-wide association studies attempt to identify genes that control variation in traits. Complex or quantitative traits are controlled by many genes (GODDARD and HAYES 2009), each contributing a proportion of the genetic variation. These traits generally have few genes that are of large or moderate effect size and many genes

with a small effect size (HAYES and GODDARD 2001). Recently, high-density single nucleotide polymorphism (SNP) panels for a broad range of species have been developed, including humans, mice, plant species – such as barley, wheat or maize – as well as major livestock species – such as cattle, pigs, sheep and chickens. Constant development of these SNP panels has lead to increased coverage of the genome, with for example as many as 1,000,000 SNPs in humans or 60,000 SNPs in pigs (GODDARD and HAYES 2009).

In 2009, the PorcineSNP60 panel, produced by the biotechnology company Illumina Inc (San Diego, California), became commercially available. This panel is based on DNA from Duroc, Landrace, Large White, Pietrain and Wild Boar individuals (RAMOS *et al.* 2009). Due to its very recent development, only a few studies have used the PorcineSNP60 panel. GORBACH *et al.* (2010a; 2010b) found no evidence for a major gene for polydactyly, but several significant SNPs associated with residual feed intake and its component traits. Other studies detected chromosomal regions and candidate genes for sow lifetime (re)production and structural soundness traits (ONTERU *et al.* 2010) and androsterone levels, a compound of boar taint (DUIJVESTEIJN *et al.* 2010a). In addition to genome-wide association studies, the PorcineSNP60 panel has also proven useful for paternal identification (DUIJVESTEIJN *et al.* 2010b), determination of inbreeding (SILIÓ *et al.* 2010) or colonisation history (SOUZA *et al.* 2010).

Behavioural traits are often difficult to measure, because their recording tends to be labour intensive and thus costly. Establishing genetic markers as indicators for behavioural traits can further enhance our understanding of the genomic basis of these traits, as for example shown in studies using a SNP panel to describe aggressive behaviour in dogs (VAGE and LINGAAS 2008) or the brain serotonergic system (TERENINA *et al.* 2010) and hypothalamic-pituitary-adrenocortical axis-related variation (MURANI *et al.* 2010) in pigs. In **Chapter 5**, analyses are described to characterise the genomic structure of a pig population based on genotypes using

the PorcineSNP60 panel, followed by **Chapter 6** in which the marker effect distribution of aggressive behaviour traits and lesion score traits is investigated.

1.3.4 Genomic selection

In the past, selective breeding in plant and livestock species was based on phenotypic information combined with extensive pedigrees and substantial progress has been achieved using this method. However, in some cases – for example when phenotypes can only be measured after several years or when traits show low heritabilities – genetic improvement per year is likely to be low.

High density SNP panels opened up the possibility of using molecular genomic information to estimate genomic breeding values. The term 'genomic selection' was coined by MEUWISSEN *et al.* (2001) for selection using genome-wide dense markers across the whole genome to estimate breeding values. The basic implementation of genomic selection consists of two steps. First, individual SNP effects are estimated in a reference or training population (MEUWISSEN 2007) to develop prediction equations of all traits under selection. Subsequently, these prediction equations are used to predict genomic breeding values of selection candidates based on their genomic SNP markers. Different strategies can be followed for the choice of reference population, for example animals with reliable breeding values, close relatives of selection candidates or animals across lines if selection candidates originate from different lines (CALUS 2009).

Estimating genomic breeding values based on SNP genotypes of an animal may provide large benefits for traits such as meat quality characteristics which can only be measured after death (MULLEN *et al.* 2006) or milk production performance, which can only be measured in female animals (SCHAEFFER 2006). Other traits that may benefit from genomic selection are characteristics in species with a large generation interval e.g. oil palm (WONG and BERNARDO 2008) or behavioural traits such as aggressiveness, which are often costly and time consuming to measure routinely (TURNER *et al.* 2009).

Initially, simulated data sets were used to assess the performance of genomic selection, evaluating the influence of different aspects such as heritability (VILLUMSEN *et al.* 2009), underlying quantitative trait loci (QTL) distribution (KIZILKAYA *et al.* 2009) or population size (MEUWISSEN 2009). The availability of SNP data in species with large numbers of recorded phenotypes offers the opportunity to evaluate the efficiency of the proposed models for genomic selection on observed data, such as mice (LEGARRA *et al.* 2008; USAI *et al.* 2009), cattle (LUAN *et al.* 2009; SU *et al.* 2009) and fish (NIELSEN *et al.* 2009). In practice, the use of genomic selection has revolutionised dairy cattle breeding (HAYES *et al.* 2009; VANRADEN and SULLIVAN 2010), but so far it has not been implemented widely in other species.

Incorporation of SNP genotypes via genomic selection for improvement of behavioural traits has the potential to benefit those costly to measure indicators of animal welfare greatly. The research described in **Chapter 4** compares the efficiency of polygenic selection, genomic selection and a combined polygenic and genomic selection approach for a range of traits in mice. In **Chapter 6**, the opportunities to incorporate genomic selection for reduction of aggressive behaviour in pigs in a breeding programme are investigated.

1.3.5 Reduction of costs of genotyping

The high cost of genotyping, especially for the high-density SNP panels, limits the extent to which routine genotyping can be implemented in practice. Additionally, many of the SNPs contribute little to the genetic variance of a trait (MEUWISSEN *et al.* 2001), as was found for example for human height variation (VISSCHER *et al.* 2007) or complex disease traits (MANOLIO *et al.* 2009). Furthermore, statistical limitations arise from the fact that many more SNP effects have to be estimated compared to the phenotypic data available. For these reasons, the ability to use fewer SNPs with a similarly high efficiency is of great interest.

Costs of genotyping may be reduced by genotyping only part of a population (e.g. OLSEN *et al.* 2010), or a two-step approach to prioritise SNPs for genotyping with low-density SNP panels (e.g. LI 2008). To circumvent the statistical limitations, many different approaches have been developed to reduce the number of effects to be estimated. One approach is the combination of SNPs into haplotypes (CALUS *et al.* 2009) which requires sophisticated methods for imputation of haplotypes or missing genotypes, for example long-range phasing (DAETWYLER *et al.* 2010; HICKEY *et al.* 2010). Another possibility is selection of subsets of SNPs (SOLBERG *et al.* 2009a; USAI *et al.* 2009; BOITARD *et al.* 2010; HARRIS and JOHNSON 2010).

In **Chapter 4**, the efficiency of genomic selection using different sizes of subsets of SNPs showing an effect is compared across a range of traits in mice. This approach is extended in **Chapter 6** to a specific set of behavioural traits to asses the efficiency of selection against aggressive behaviour using reduced quantities of SNPs.

1.4 Thesis aims

As described in the previous sections, improvement of animal welfare traits could have substantial benefits to producers as well as consumers, and is therefore of increasing importance in our existing production systems. The overall aim of this thesis was to investigate if and how genetic improvement of characteristics associated with animal welfare can be achieved using quantitative and molecular genetic approaches. The specific objectives of the thesis were as follows:

- To estimate genetic parameters of piglet survival at birth and its associations with piglet birth weight, at litter level and at piglet level, in order to examine whether additional selection for birth weight can contribute to improvement of survival at birth over and above direct selection for survival.
- To estimate genetic associations between piglet survival at birth and economically important reproduction traits, such as number born alive and

number born in total, and production traits, such as average daily gain, backfat thickness and muscle depth.

- To investigate differences in genetic parameters between sire and dam lines, originating from different breeds or from the same breed but separated 25 years ago and selected based on different breeding goals, and change in genetics parameters and associations between traits due to selection.
- To investigate the influence of SNP subset size, heritability, QTL-distribution and type of trait (behavioural versus physiological) on the efficiency of genomic selection.
- To establish the genomic structure of a pig population based on genotypes, including chromosome coverage, minor allele frequencies, Hardy-Weinberg Equilibrium and linkage disequilibrium.
- To evaluate the efficiency of genomic selection against aggressive behaviour in pigs compared to polygenic selection or a combined genomic and polygenic approach.

1.5 Thesis outline

This thesis consists of six further chapters, whose general objectives are described as follows: Chapter 2 discusses the associations between the traits piglet survival and birth weight and the opportunities to improve piglet survival using selection for birth weight as an indirect trait. Chapter 3 explores the genetic correlations between piglet survival and economically important traits such as backfat and litter size and the change in genetic parameters due to selection. In Chapter 4, the performance of a methodology to incorporate molecular genetic information into the selection process is assessed and the efficiency of genomic selection on a range of traits in mice is discussed. Chapter 5 describes the genomic structure of a data set which included molecular genetic information of a high-density SNP panel in pigs. In Chapter 6 the methodology examined in Chapter 4 is used on the genomic data described in Chapter 5 to evaluate the efficiency of genomic selection against aggressive behaviour traits in pigs. The final chapter is a general discussion in which findings

from all chapters of the thesis are discussed and final conclusions are derived for the analysed traits associated with animal welfare. Moreover, suggestions for future research in this area are given based on the obtained results. A combined bibliography for all chapters is included at the end of this thesis.

Chapter 2 – Improving piglet survival via selection for birth weight

2.1 Introduction

The aim of this chapter was to estimate heritabilities and genetic correlations of survival and weight traits at birth, both at litter level as well as at individual piglet level, for different breeds. These genetic parameters give insight into the efficiency of selection for survival, thus allowing the examination of whether additional selection for piglet birth weight can contribute to improvement of piglet survival at birth over and above direct selection for survival.

2.2 Materials and methods

2.2.1 Animals

Data on piglet survival, litter size and individual birth weight in several sire and dam lines were provided by the pig breeding organisation PIC. Data were available for 36,217 purebred piglets born out of 3497 litters between January 2005 and September 2006 on two farms in China and one in Brazil. One farm, in China, was located in a northerly continental, monsoonal climate; the two remaining farms, in China and Brazil, in a humid subtropical climate. Data included birth weight measured individually on piglets born alive, and for each litter the numbers of piglets born in total, born alive and stillborn was known. Information regarding crossfostering or litters with no piglets born alive was not available. The piglets came from five different purebred lines: dam line D1 (955 litters, average 1.4 litters per dam), dam line D2 (1302 litters, average 1.4 litters per dam), dam line D3 (504 litters, average 1.7 litters per dam), sire line S1 (253 litters, average 1.2 litters per dam) and sire line S2 (483 litters, average 1.3 litters per dam). The three dam lines were bred with the main emphasis on litter size and piglet survival, with additional emphasis on growth rate in D3. Sire line S1 was bred with emphasis on body leanness; sire line S2 with emphasis on growth rate, feed efficiency and growing pig survival.

Four traits were analysed at litter level: percentage survival at birth (SVBL), number born in total (NBT), average birth weight of piglets within litter (ALBW) and standard deviation of birth weight within litter (STD). At individual piglet level, two traits were analysed: individual piglet survival at birth (SVBP) and individual piglet birth weight (IBW). SVBL was calculated as the number of piglets born alive divided by the number of piglets born in total. ALBW was calculated as the sum of all birth weights per litter divided by the number of piglets with known birth weight (stillborn piglets and piglets with missing birth weight were not considered). For STD only litters with a minimum of five piglets with known birth weight were included. It should be noted that litter size and standard deviation of piglet birth weight are to some extent correlated. For a given size of the uterus, smaller litters generally have larger average piglet weights and standard deviations of piglet weight, whereas larger litters with lower average weights show lower standard deviations. Most traits were approximately normally distributed; percentage survival a birth at litter level showed a slightly skewed distribution while survival at piglet level was binomially distributed, but transformation was not considered necessary. Pre-weaning survival after farrowing was not included in this analysis due to very high piglet survival levels (98% on average) on these farms, which led to very low levels of detectable genetic variation in the data. The quality of pedigree files was checked with Relax2 (STRANDÉN and VUORI 2006). No errors in the pedigree were found, after which pedigree files were matched per line to the animals in the data set to eliminate redundant animals. No limit was set for the depth of the pedigree and pedigrees consisted of 2929 (D1), 3326 (D2), 1591 (D3), 566 (S1) and 1406 (S2) animals at litter level.

2.2.2 Analysis at litter level

The traits SVBL, NBT, ALBW and STD were analysed for all five lines separately. In addition, to increase the number of observations and decrease the confidence interval of estimates, the two dam lines D1 and D2, with the same breeding goals, were analysed jointly as D12. In a preliminary analysis, fixed effects were tested per trait and line for significance using the procedures MIXED or GENMOD (SAS

2002). Based on these analyses parity, gestation length and batch were included in the subsequent models, analysed using a Bayesian approach. Batches were optimised in preliminary analyses and fitted in the model based on farm, year and quarter of year (starting with January to March). In the dam lines, parities one to five were considered as separate classes, parities six and higher were grouped together. In D12, parities one to seven were considered separately and parities eight and higher were grouped together. In the sire lines, only two parity-classes were significant: either first parity or higher parities. Gestation length was grouped as ≤ 113, 114, ..., 117, ≥ 118 for D1 and D2, ≤ 114, 115, ..., 117, ≥ 118 for D3, ≤ 112, 113, 117, ≥ 118 for D12, ≤ 113, 114, 115, ≥ 116 for S1 and ≤ 115 or ≥ 116 for S2. All effects were grouped such that sufficiently high numbers of observations (at least 40 observations) were available for each group. Approximately 65% to 70% of the dams in lines D1, D2 and S2 had only one litter in the data set, while the remaining dams produced up to four litters. For S1 more than 83% of the dams produced only one litter, while for D3 only 55% of the dams produced one litter and 45% of the dams produced up to four litters. The multiple trait model to estimate genetic parameters was as follows:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{Wc} + \boldsymbol{e},$$

where $\boldsymbol{y}$ is the vector of observations of the traits, $\boldsymbol{b}$ the vector of fixed effects of parity, gestation length and batch, $\boldsymbol{a}$ the vector of additive genetic effects, $\mathbf{c}$ the vector of permanent environmental effects of the dam and $\boldsymbol{e}$ the vector of residuals. $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are incidence matrices relating the vectors $\boldsymbol{b}$, $\boldsymbol{a}$, and $\boldsymbol{c}$ with $\boldsymbol{y}$. The assumed (co)variance structure was:

$$V\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{c}\\ \boldsymbol{e}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}\otimes\boldsymbol{G} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\otimes\boldsymbol{C} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\otimes\boldsymbol{R}\end{bmatrix},$$

where ***A*** and ***I*** are the additive genetic relationship matrix and identity matrix, respectively. ***G, C*** and ***R*** represent the variance and covariance matrices of additive genetic effects, permanent environmental effects of the dam and residual effects, respectively.

2.2.3 Analysis at piglet level

Preliminary analyses of SVBP and IBW indicated that no convergence of parameters could be achieved in the multiple trait analyses of SVBP and IBW of lines D3, S1 and S2. This is possibly due to the much lower number of observations in these lines compared to D1 and D2. Thus, only D1 and D2 were analysed at piglet level, once separately for each line and then in a combined analysis of both lines, D12, to increase the number of observations and precision of estimates.

Fixed effects were tested per trait and line for significance using the procedures MIXED or GENMOD (SAS 2002) and included parity, gestation length and batch for both traits. For IBW, sex of the piglet was also significant and included in the model. Batches were optimised in preliminary analyses and fitted based on farm, year and quarter of year (starting with January to March). Parities for D1 (D2 and D12) were grouped in nine (ten) classes with the first eight (nine) parities as separate classes and higher parities combined. Gestation length was grouped as $\leq 112, 113, ..., 119, \geq 120$ for D1, as $\leq 111, 112, ..., 118, \geq 119$ for D2 and as $\leq 111, 112, ..., 119, \geq 120$ for D12. Effects were grouped such that at least 40 observations were available per group. In a further analysis both traits were adjusted for litter size. Litter size (number of piglets born in total) was included as a categorical effect with groups of $\leq 3, 4, ..., 18, \geq 19$ piglets per litter for D1 and D2 and $\leq 2, 3, ..., 19, \geq 20$ for D12. SVBP, scored as zero (dead) or one (alive), was analysed using a threshold model. The trait IBW was normally distributed, while SVBP followed a binomial distribution, therefore the multiple trait threshold model used to estimate genetic parameters using a Bayesian approach was as follows:

$$\boldsymbol{y = Xb + Z_1d + Z_2m + Wl + e,}$$

where ***y*** is the vector of observations of the traits, ***b*** the vector of fixed effects of parity, gestation length and batch fitted for both traits and additionally the sex effect fitted for IBW, ***d*** the vector of direct genetic effects, ***m*** the vector of maternal genetic effects, ***l*** the vector of common environmental effects of the litter and **e** the vector of residuals. The vector **y** contains the observed phenotype for IBW as well as an unobservable continuous variable for SVBP, relating the observed phenotype through a threshold. $\boldsymbol{X}$, $\boldsymbol{Z_1}$, $\boldsymbol{Z_2}$ and $\boldsymbol{W}$ are incidence matrices relating the vectors ***b***, ***d***, ***m*** and ***l*** with ***y***. The assumed (co)variance structure was:

$$V\begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{m} \\ \boldsymbol{l} \\ \boldsymbol{e} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}\otimes\boldsymbol{G}_d & \boldsymbol{A}\otimes\boldsymbol{G}_{dm} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{A}\otimes\boldsymbol{G}_{md} & \boldsymbol{A}\otimes\boldsymbol{G}_m & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\otimes\boldsymbol{L} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\otimes\boldsymbol{R} \end{bmatrix},$$

where ***A*** and ***I*** are the additive genetic relationship matrix and identity matrix, respectively. $\boldsymbol{G_d}$, $\boldsymbol{G_m}$, $\boldsymbol{G_{dm}}$, ***L*** and ***R*** represent the variance and covariance matrices of direct genetic effects, maternal genetic effects, matrix of covariances between direct and maternal genetic effects, variance and covariance matrices of common environmental effects of the litter and residual effects, respectively. The residual variance for the fitted threshold model was set to 1 and other variances were estimated relative to the residual variance. Direct (σ^2_d), maternal (σ^2_m), covariance between the direct and maternal genetic (σ_{dm}), litter (σ^2_l), residual (σ^2_e) and phenotypic variances $(\sigma^2_p = \sigma^2_d + \sigma_{dm} + \sigma^2_m + \sigma^2_l + \sigma^2_e$ (WILLHAM 1972)) were estimated for both traits. Direct (h^2_d), maternal (h^2_m) and total heritability $(h^2_t = (\sigma^2_d + 0.5 * \sigma^2_m + 1.5 * \sigma_{dm}) / \sigma^2_p$ (WILLHAM 1972)) were calculated.

2.2.4 Bayesian analysis

Data were analysed based on linear models (litter level) or combined linear-threshold models (piglet level). A Bayesian approach was used, using Gibbs sampling and MCMC methods with flat priors as implemented in the software of GIBBS2F90 and THRGIBBS1F90, respectively (MISZTAL *et al.* 2002). Bayesian confidence intervals

were calculated to determine the precision of the estimates. Posterior means of parameters were calculated based on the marginal posterior distributions obtained using Gibbs sampling. Convergence of the chains was assessed using the Geweke criterion (GEWEKE 1992), Raftery and Lewis criterion (RAFTERY and LEWIS 1992) and visual assessment of the drawn marginal posterior distributions.

All estimated variance components, heritabilities and correlations of traits were calculated as the mean of the marginal posterior distribution; the precision of these estimates was determined based on the 95%-Highest Posterior Density interval (95%-HPD). Depending on line and traits, chains of up to 1,000,000 iterations were used, with up to 300,000 iterations at the start of the chain discarded. To reduce the autocorrelation between the saved samples, a lag of 50 was used, so that only every 50^{th} iteration was used to estimate the marginal posterior distribution.

2.3 Results

2.3.1 Descriptive results

The mean percentage of piglets that survived at birth ranged from 92.1% to 94.6%. The three dam lines, selected with emphasis on survival, showed the highest percentages of survival (Tables 2.1 and 2.2). D1 had the largest litter size (10.7 piglets) with on average 0.6 piglets more born than the line with the lowest litter size (S2). Differences in both ALBW and IBW were significant between all lines ($P < 0.05$) except between D1 and D3. S2, selected for growth rate, feed efficiency and growing pig survival, showed the highest average birth weights, while S1, selected for body leanness, showed the lowest. Average birth weight was approximately 10% to 25% higher for S2 compared to the other four lines.

Table 2.1: Summary statistics at litter level. Descriptive statistics of the traits and lines (N = number of litters, s.e. as subscript of the mean, CV = coefficient of variation). Means with different superscripts within column are significantly different (P < 0.05 with Bonferroni correction to account for multiple means comparison).

Dam line	N	SVBL (%)	NBT	ALBW (g)	STD (g)	CV (%)
D1	955	$93.3_{0.35}{}^{a}$	$10.7_{0.11}{}^{a}$	$1452_{7.9}{}^{b}$	$250_{2.9}{}^{a}$	17.2
D2	1302	$93.5_{0.30}{}^{a}$	$10.2_{0.09}{}^{b}$	$1386_{6.8}{}^{c}$	$222_{2.5}{}^{b}$	16.0
D3	504	$94.6_{0.48}{}^{a}$	$10.1_{0.15}{}^{b}$	$1443_{10.9}{}^{b}$	$242_{4.1}{}^{a}$	16.8
Sire line						
S1	253	$92.1_{0.68}{}^{a}$	$10.6_{0.21}{}^{ab}$	$1277_{15.3}{}^{d}$	$209_{5.8}{}^{b}$	16.4
S2	483	$92.4_{0.49}{}^{a}$	$10.1_{0.15}{}^{b}$	$1590_{11.1}{}^{a}$	$251_{4.2}{}^{a}$	15.8

S2 had the highest variation of birth weight within litter, but did not differ significantly from D1 and D3 (250 g and 242 g). However, these lines had significantly higher variation than D2 and S1 ($P < 0.05$; 222 g and 209 g, respectively).

Table 2.2: Summary statistics at piglet level. Descriptive statistics of the traits and lines (N = number of piglets, s.e. as subscript of the mean). Means with different superscripts within column are significantly different (P < 0.05 with Bonferroni correction to account for multiple means comparison).

Dam line	N	SVBP (%)	IBW (g)
D1	10,259	93.4	$1420_{3.4}{}^{b}$
D2	13,306	93.5	$1358_{2.9}{}^{c}$
D3	5095	94.5	$1421_{4.7}{}^{b}$
Sire line			
S1	2692	92.7	$1258_{6.6}{}^{d}$
S2	4865	92.7	$1547_{4.9}{}^{a}$

Table 2.3: Extreme weights. Percentage of litters with piglets weighing outside two standard deviations of the mean of the ALBW (upper and lower weight).

Dam line	Lower weight (g)	Upper weight (g)	Low (%)	High (%)
D1	947	1957	18.6	5.7
D2	903	1869	15.6	8.4
D3	976	1909	15.3	6.6
Sire line				
S1	835	1720	14.2	12.4
S2	1088	2092	21.1	4.5

Table 2.3 shows that S2 had a higher percentage (21%) of litters with runts (piglets weighing less than ALBW minus two times the standard deviation within line) than all other lines (14% to 19%). Conversely, S2 had the lowest percentage (5%) of litters with exceptionally heavy piglets (a piglet weighing more than ALBW plus two times the standard deviation within line).

2.3.2 Litter level heritabilities

The posterior means of phenotypic variances, heritabilities and phenotypic proportion of the permanent environmental effect of litter traits are presented in Table 2.4. Posterior means of the phenotypic variance for SVBL varied considerably among the lines, with S1 showing a substantially higher heritability of 0.20 for survival, compared to 0.05 to 0.13 in the other lines.

NBT showed similar heritabilities among lines, in the range of 0.11 to 0.16. Phenotypic variances and heritabilities for ALBW were lower in the three dam lines and S1 than in S2. The range of heritabilities for ALBW was higher than for NBT, from 0.23 in D1 up to 0.34 in S2. For STD, sire line S1 showed the highest heritability (0.27) of all lines. The combined analysis of dam lines, D12, showed posterior means of estimates that were generally between the values estimated for the separate lines, but with a smaller 95%-HPD and thus much more precise.

Table 2.4: Variance components and heritabilities at litter level. Posterior means of phenotypic variance (σ^2_p), heritability (h^2) and phenotypic proportion of the permanent environmental effect (PE) in dam and sire lines (95%-HPD as subscript).

		Dam Line				Sire Line	
Trait		**D1**	**D2**	**D3**	**D12**	**S1**	**S2**
SVBL	σ^2_p	93.59 84.89 to 102.14	110.20 101.17 to 119.6	75.48 65.91 to 85.82	102.7 96.43 to 109.08	179.10 145.36 to 216.32	135.00 116.70 to 152.75
	h^2	0.05 0.01 to 0.12	0.13 0.03 to 0.23	0.08 0.00 to 0.17	0.11 0.04 to 0.18	0.20 0.01 to 0.38	0.08 0.00 to 0.18
	PE	0.08 0.00 to 0.19	0.08 0.01 to 0.16	0.06 0.00 to 0.13	0.08 0.01 to 0.15	0.12 0.00 to 0.27	0.10 0.01 to 0.21
NBT	σ^2_p	10.83 9.83 to 11.93	11.32 10.41 to 12.26	8.88 7.68 to 10.11	10.95 10.26 to 11.63	9.95 8.08 to 11.93	10.89 9.44 to 12.36
	h^2	0.16 0.06 to 0.26	0.16 0.09 to 0.25	0.11 0.01 to 0.21	0.16 0.09 to 0.22	0.13 0.01 to 0.27	0.11 0.02 to 0.22
	PE	0.09 0.01 to 0.18	0.08 0.02 to 0.16	0.13 0.02 to 0.24	0.07 0.01 to 0.13	0.14 0.02 to 0.28	0.10 0.01 to 0.20
ALBW	σ^2_p	56,750 51,140 to 62,810	53,460 48,820 to 58,472	57,940 49,558 to 67,387	51,510 47,828 to 55,231	55,470 44,572 to 67,226	66,860 57,746 to 76,950
	h^2	0.23 0.12 to 0.35	0.30 0.20 to 0.41	0.26 0.09 to 0.42	0.27 0.19 to 0.36	0.26 0.08 to 0.45	0.34 0.21 to 0.48
	PE	0.21 0.09 to 0.33	0.17 0.07 to 0.27	0.15 0.03 to 0.28	0.15 0.06 to 0.24	0.19 0.04 to 0.35	0.12 0.03 to 0.23
STD	σ^2_p	8342.0 7539.4 to 9178.0	7094.0 6473.6 to 7703.7	9058.0 7797.6 to 10,418.0	7458.0 6969.2 to 7910.1	7600.0 5973.3 to 9342.0	9594.0 8259.9 to 11,018.0
	h^2	0.11 0.04 to 0.18	0.10 0.03 to 0.17	0.10 0.02 to 0.19	0.13 0.08 to 0.18	0.27 0.10 to 0.46	0.14 0.05 to 0.26
	PE	0.09 0.01 to 0.16	0.14 0.06 to 0.24	0.14 0.04 to 0.24	0.07 0.02 to 0.13	0.21 0.06 to 0.38	0.13 0.03 to 0.24

2.3.3 Litter level correlations

Tables 2.5 and 2.6 show the posterior means of the genetic, permanent environmental, residual and phenotypic correlations between litter traits for each of the five individual lines and the two dam lines combined. Genetic correlations between SVBL and NBT were, with a probability of $Pr(r_g \leq 0) = 0.62$ to 0.78, negative or zero and thus unfavourable in D3 and D2, respectively, but positive and thus favourable, with a probability of $Pr(r_g > 0) = 0.62$ to 0.85, in the other lines. More consistent among lines was the genetic correlation between SVBL and ALBW, which was positive with a probability of $Pr(r_g > 0) = 0.73$ to 0.93 in all lines except D1.

Table 2.5: Correlations at litter level in the dam lines. Posterior means of genetic (r_g), permanent environmental (r_{pe}), residual (r_e) and phenotypic (r_p) correlations between litter traits (95%-HPD and posterior probability of a positive correlation Pr(r>0) as subscript, estimates of correlations with 95%-HPD excluding zero are bold) (continued on next page).

Trait		D1	D2	D3	D12
SVBL - NBT	r_g	0.39 -0.21 to 1.00 / Pr(rg>0) = 0.85	-0.22 -0.80 to 0.28 / Pr(rg>0) = 0.22	-0.12 -0.88 to 0.65 / Pr(rg>0) = 0.38	0.18 -0.21 to 0.53 / Pr(rg>0) = 0.83
	r_{pe}	0.04 -0.89 to 0.84 / Pr(rpe>0) = 0.55	0.39 -0.31 to 1.00 / Pr(rpe>0) = 0.84	-0.30 -0.99 to 0.54 / Pr(rpe>0) = 0.23	-0.07 -0.77 to 0.62 / Pr(rpe>0) = 0.42
	r_e	-0.04 -0.15 to 0.06 / Pr(re>0) = 0.24	-0.02 -0.11 to 0.06 / Pr(re>0) = 0.31	-0.04 -0.15 to 0.08 / Pr(re>0) = 0.25	-0.02 -0.09 to 0.04 / Pr(re>0) = 0.27
	r_p	0.01 -0.06 to 0.07 / Pr(rp>0) = 0.58	-0.01 -0.07 to 0.05 / Pr(rp>0) = 0.35	-0.07 -0.16 to 0.02 / Pr(rp>0) = 0.07	0.00 -0.04 to 0.05 / Pr(rp>0) = 0.54
SVBL – ALBW	r_g	-0.50 -1.00 to 0.18 / Pr(rg>0) = 0.09	0.35 -0.14 to 0.90 / Pr(rg>0) = 0.92	0.55 -0.09 to 0.99 / Pr(rg>0) = 0.93	0.06 -0.31 to 0.44 / Pr(rg>0) = 0.62
	r_{pe}	0.07 -0.76 to 0.80 / Pr(rpe>0) = 0.58	-0.54 -1.00 to 0.07 / Pr(rpe>0) = 0.07	0.46 -0.48 to 1.00 / Pr(rpe>0) = 0.84	-0.20 -0.73 to 0.38 / Pr(rpe>0) = 0.22
	r_e	0.09 -0.02 to 0.20 / Pr(re>0) = 0.94	**0.15 0.06 to 0.24 / Pr(re>0) = 1.00**	-0.01 -0.13 to 0.10 / Pr(re>0) = 0.42	**0.13 0.05 to 0.21 / Pr(re>0) = 1.00**
	r_p	0.02 -0.05 to 0.09 / Pr(rp>0) = 0.73	**0.10 0.04 to 0.16 / Pr(rp>0) = 1.00**	**0.12 0.02 to 0.21 / Pr(rp>0) = 0.99**	**0.08 0.02 to 0.13 / Pr(rp>0) = 1.00**

Trait		D1	D2	D3	D12
SVBL - STD	r_g	-0.52[a] -1.00 to 0.27 / Pr(rg>0) = 0.12	-0.18[a] -0.86 to 0.80 / Pr(rg>0) = 0.32	-0.18 -0.99 to 0.60 / Pr(rg>0) = 0.34	-0.39 -0.75 to 0.01 / Pr(rg>0) = 0.03
	r_{pe}	0.29 -0.70 to 1.00 / Pr(rpe>0) = 0.70	0.25 -0.61 to 0.99 / Pr(rpe>0) = 0.74	-0.03 -0.91 to 0.89 / Pr(rpe>0) = 0.47	**0.69 0.18 to 1.00 / Pr(rpe>0) = 0.98**
	r_e	-0.04 -0.15 to 0.08 / Pr(re>0) = 0.28	-0.03 -0.14 to 0.07 / Pr(re>0) = 0.26	0.07 -0.06 to -0.21 / Pr(re>0) = 0.86	-0.05 -0.12 to 0.03 / Pr(re>0) = 0.11
	r_p	-0.04 -0.13 to 0.04 / Pr(rp>0) = 0.16	-0.02 -0.10 to 0.06 / Pr(rp>0) = 0.33	0.04 -0.08 to 0.15 / Pr(rp>0) = 0.76	-0.03 -0.09 to 0.03 / Pr(rp>0) = 0.15
NBT - ALBW	r_g	**-0.40 -0.78 to -0.01 / Pr(rg>0) = 0.03**	**-0.43 -0.70 to -0.15 / Pr(rg>0) = 0.01**	-0.48 -0.99 to 0.07 / Pr(rg>0) = 0.07	**-0.39 -0.61 to -0.15 / Pr(rg>0) = 0.00**
	r_{pe}	-0.21 -0.80 to 0.45 / Pr(rpe>0) = 0.22	**-0.63 -0.97 to -0.23 / Pr(rpe>0) = 0.01**	-0.30 -0.90 to 0.36 / Pr(rpe>0) = 0.17	**-0.54 -0.91 to -0.11 / Pr(rpe>0) = 0.02**
	r_e	**-0.47 -0.55 to -0.38 / Pr(re>0) = 0.00**	**-0.39 -0.47 to -0.32 / Pr(re>0) = 0.00**	**-0.28 -0.40 to -0.17 / Pr(re>0) = 0.00**	**-0.46 -0.52 to -0.41 / Pr(re>0) = 0.00**
	r_p	**-0.42 -0.47 to -0.36 / Pr(rp>0) = 0.00**	**-0.42 -0.47 to -0.37 / Pr(rp>0) = 0.00**	**-0.31 -0.41 to -0.22 / Pr(rp>0) = 0.00**	**-0.45 -0.49 to -0.41 / Pr(rp>0) = 0.00**
NBT - STD	r_g	-0.30 -0.79 to 0.20 / Pr(rg>0) = 0.14	**0.46 0.02 to 0.83 / Pr(rg>0) = 0.97**	-0.13 -0.90 to 0.61 / Pr(rg>0) = 0.38	0.16 -0.15 to 0.52 / Pr(rg>0) = 0.82
	r_{pe}	-0.05 -0.75 to 0.71 / Pr(rpe>0) = 0.46	-0.22 -0.81 to 0.37 / Pr(rpe>0) = 0.24	0.23 -0.44 to 0.88 / Pr(rpe>0) = 0.75	-0.34 -0.99 to 0.21 / Pr(rpe>0) = 0.15
	r_e	**0.15 0.06 to 0.26 / Pr(re>0) = 1.00**	**0.21 0.12 to 0.30 / Pr(re>0) = 1.00**	**0.19 0.07 to 0.32 / Pr(re>0) = 1.00**	**0.18 0.11 to 0.24 / Pr(re>0) = 1.00**
	r_p	0.08 0.00 to 0.15 / Pr(rp>0) = 0.97	**0.19 0.13 to 0.26 / Pr(rp>0) = 1.00**	**0.17 0.05 to 0.28 / Pr(rp>0) = 1.00**	**0.14 0.09 to 0.19 / Pr(rp>0) = 1.00**
ALBW - STD	r_g	**0.58 0.21 to 0.92 / Pr(rg>0) = 1.00**	0.29 -0.14 to 0.73 / Pr(rg>0) = 0.90	0.29 -0.36 to 0.86 / Pr(rg>0) = 0.82	**0.47 0.20 to 0.72 / Pr(rg>0) = 1.00**
	r_{pe}	0.33 -0.18 to 0.82 / Pr(rpe>0) = 0.89	0.30 -0.17 to 0.76 / Pr(rpe>0) = 0.89	0.37 -0.20 to 0.90 / Pr(rpe>0) = 0.89	0.29 -0.23 to 0.75 / Pr(rpe>0) = 0.87
	r_e	0.00 -0.11 to 0.10 / Pr(re>0) = 0.49	-0.09 -0.18 to 0.01 / Pr(re>0) = 0.04	0.00 -0.12 to 0.13 / Pr(re>0) = 0.52	-0.06 -0.13 to 0.01 / Pr(re>0) = 0.03
	r_p	**0.13 0.06 to 0.21 / Pr(rp>0) = 1.00**	0.04 -0.02 to 0.10 / Pr(rp>0) = 0.89	0.11 0.00 to 0.22 / Pr(rp>0) = 0.97	**0.07 0.03 to 0.12 / Pr(rp>0) = 1.00**

[a] no convergence

Table 2.6: Correlations at litter level in the sire lines. Posterior means of genetic (r_g), permanent environmental (r_{pe}), residual (r_e) and phenotypic (r_p) correlations between litter traits (95%-HPD and posterior probability of a positive correlation Pr(r>0) as subscript, estimates of correlations with 95%-HPD excluding zero are bold).

Trait		S1	S2
SVBL - NBT	r_g	0.14 -0.72 to 0.99 / Pr(rg>0) = 0.62	0.31 -0.44 to 1.00 / Pr(rg>0) = 0.79
	r_{pe}	0.49 -0.36 to 1.00 / Pr(rpe>0) = 0.87	0.10 -0.75 to 0.90 / Pr(rpe>0) = 0.61
	r_e	0.03 -0.16 to 0.22 / Pr(re>0) = 0.63	0.08 -0.05 to 0.21 / Pr(re>0) = 0.88
	r_p	0.12 -0.02 to 0.25 / Pr(rp>0) = 0.95	**0.11 0.02 to 0.20 / Pr(rp>0) = 0.99**
SVBL - ALBW	r_g	0.47 -0.25 to 1.00 / Pr(rg>0) = 0.89	0.22 -0.43 to 0.86 / Pr(rg>0) = 0.73
	r_{pe}	0.14[a] -0.80 to 1.00 / Pr(rpe>0) = 0.61	-0.58 -1.00 to 0.44 / Pr(rpe>0) = 0.12
	r_e	0.02 -0.19 to 0.23 / Pr(re>0) = 0.59	-0.09 -0.24 to 0.06 / Pr(re>0) = 0.13
	r_p	0.15 0.00 to 0.28 / Pr(rp>0) = 0.98	-0.10 -0.19 to 0.00 / Pr(rp>0) = 0.03
SVBL - STD	r_g	-0.43 -1.00 to 0.34 / Pr(rg>0) = 0.15	0.48 -0.33 to 1.00 / Pr(rg>0) = 0.86
	r_{pe}	-0.22[a] -1.00 to 0.71 / Pr(rpe>0) = 0.33	0.04 -0.82 to 0.97 / Pr(rpe>0) = 0.53
	r_e	**0.33 0.05 to 0.58 / Pr(re>0) = 0.99**	-0.01 -0.17 to 0.15 / Pr(re>0) = 0.45
	r_p	0.07 -0.15 to 0.30 / Pr(rp>0) = 0.73	0.05 -0.08 to 0.18 / Pr(rp>0) = 0.77
NBT - ALBW	r_g	-0.01 -0.76 to 0.75 / Pr(rg>0) = 0.49	**-0.54 -0.91 to -0.14 / Pr(rg>0) = 0.01**
	r_{pe}	-0.03 -0.76 to 0.77 / Pr(rpe>0) = 0.46	-0.15 -0.81 to 0.61 / Pr(rpe>0) = 0.33
	r_e	**-0.63 -0.76 to -0.48 / Pr(re>0) = 0.00**	**-0.64 -0.73 to -0.55 / Pr(re>0) = 0.00**
	r_p	**-0.41 -0.53 to -0.29 / Pr(rp>0) = 0.00**	**-0.54 -0.61 to -0.47 / Pr(rp>0) = 0.00**
NBT - STD	r_g	0.08 -0.73 to 0.83 / Pr(rg>0) = 0.59	0.40 -0.24 to 0.93 / Pr(rg>0) = 0.88
	r_{pe}	-0.07 -0.86 to 0.70 / Pr(rpe>0) = 0.44	-0.03 -0.78 to 0.75 / Pr(rpe>0) =0.47
	r_e	0.14 -0.10 to 0.39 / Pr(re>0) = 0.86	**0.21 0.07 to 0.36 / Pr(re>0) = 1.00**
	r_p	0.09 -0.08 to 0.26 / Pr(rp>0) = 0.85	**0.21 0.09 to 0.32 / Pr(rp>0) = 1.00**
ALBW - STD	r_g	0.11 -0.53 to 0.71 / Pr(rg>0) = 0.64	0.03 -0.52 to 0.56 / Pr(rg>0) = 0.56
	r_{pe}	0.27 -0.38 to 0.86 / Pr(rpe>0) = 0.79	-0.07 -0.75 to 0.63 / Pr(rpe>0) = 0.43
	r_e	0.06 -0.18 to 0.29 / Pr(re>0) = 0.69	-0.10 -0.26 to 0.04 / Pr(re>0) = 0.10
	r_p	0.12 -0.04 to 0.28 / Pr(rp>0) = 0.93	-0.06 -0.18 to 0.04 / Pr(rp>0) = 0.13

[a] no convergence

The posterior means of genetic correlations between SVBL and STD were negative in four lines and their probabilities to be zero or negative ranged from $Pr(r_g \leq 0) = 0.66$ to 0.88. A negative correlation between SVBL and STD indicates that less variation in birth weight within litter results in higher piglet survival. Phenotypic correlations between SVBL and NBT, ALBW or STD were low and for S2 with a probability of $Pr(r_p > 0) = 0.99$ positive between SVBL and NBT and for D2 and D3 with a probability of $Pr(r_p > 0) = 0.99$ positive between SVBL and ALBW.

The posterior means of the genetic correlations between NBT and ALBW were generally moderately negative, ranging from -0.39 to -0.54 (except for S1, which was close to zero), and their probabilities to be positive were very low in the range of $Pr(r_g > 0) = 0.00$ to 0.07 (except for S1, which was $Pr(r_g > 0) = 0.49$). The genetic correlations between NBT and STD showed a diverse pattern, ranging from moderately negative in D1 (-0.30) to positive in D2 (0.46). Genetic correlations between ALBW and STD were, with probabilities of $Pr(r_g > 0) = 0.56$ to 1.00, positive and unfavourable for all lines, with posterior means ranging from 0.03 to 0.58.

2.3.4 Piglet level heritabilities

Posterior means of phenotypic variance and genetic and environmental parameters of SVBP and IBW are presented in Table 2.7. In this section the genetic parameters unadjusted for litter size in lines D1 and D2 are described. The combined analysis D12 unadjusted for litter size was not presented in Table 2.7 since the heritability of direct genetic effect for SVBP approached zero.

Direct heritabilities were very small at 0.01 and the 95%-HPD always included zero. Maternal heritabilities were substantially higher than direct heritabilities at 0.06 to 0.07 for line D1 and D2, respectively, and the lower levels of their 95%-HPD were at least 0.02. The genetic correlations between the direct and maternal effects were positive, with a probability of $Pr(r_g > 0) = 0.75$ and 1.00 for lines D1 and D2, respectively, and total heritability was 0.04 for line D1 and 0.06 for line D2.

Table 2.7: Variance components and heritabilities at piglet level. Posterior means of total phenotypic variance (σ^2_p), direct (h^2_d), maternal (h^2_m) and total heritability (h^2_t), correlation between direct and maternal genetic effect ($r_{g\text{-}dm}$) and phenotypic proportion of the litter effect (LE) for piglet traits in dam lines D1, D2 and D12 unadjusted (UAJ) and adjusted (AJ) for litter size (95%-HPD and posterior probability of a positive correlation $Pr(r_g>0)$ as subscript; estimates of correlations with 95%-HPD excluding zero bold).

		D1		**D2**		**D12**
Trait		UAJ	AJ	UAJ	AJ	AJ
SVBP	σ^2_p	1.120	1.135	1.172	1.197	1.149
		1.058 to 1.181	1.071 to 1.201	1.113 to 1.234	1.125 to 1.272	1.106 to 1.193
	h^2_d	0.01	0.02	0.01	0.03	0.01
		0.00 to 0.01	0.01 to 0.04	0.00 to 0.01	0.00 to 0.07	0.00 to 0.02
	h^2_m	0.06	0.08	0.07	0.07	0.08
		0.02 to 0.10	0.02 to 0.13	0.03 to 0.10	0.03 to 0.12	0.04 to 0.12
	h^2_t	0.04	0.04	0.06	0.06	0.02
		0.01 to 0.07	0.01 to 0.06	0.03 to 0.09	0.01 to 0.11	0.01 to 0.04
	$r_{g\text{-}dm}$	0.15	-0.41	**0.73**	-0.17	**-0.62**
		-0.30 to 0.54	-0.85 to 0.03	**0.53 to 0.89**	-0.77 to 0.43	**-0.90 to -0.27**
		Pr(rg>0) = 0.75	Pr(rg>0) = 0.05	**Pr(rg>0) = 1.00**	Pr(rg>0) = 0.30	**Pr(rg>0) = 0.01**
	LE	0.04	0.04	0.06	0.06	0.06
		0.00 to 0.08	0.00 to 0.08	0.02 to 0.10	0.02 to 0.11	0.02 to 0.09
IBW	σ^2_p	0.130	0.120	0.112	0.100	0.108
		0.122 to 0.137	0.113 to 0.126	0.106 to 0.119	0.096 to 0.105	0.104 to 0.112
	h^2_d	0.13	0.09	0.19	0.03	0.06
		0.04 to 0.25	0.02 to 0.17	0.09 to 0.30	0.01 to 0.06	0.02 to 0.11
	h^2_m	0.28	0.22	0.16	0.21	0.17
		0.16 to 0.39	0.12 to 0.31	0.08 to 0.25	0.14 to 0.28	0.12 to 0.22
	h^2_t	0.08	0.10	0.28	0.11	0.14
		0.01-0.16	0.02 – 0.18	0.17-0.39	0.06 – 0.17	0.08 to 0.19
	$r_{g\text{-}dm}$	**-0.64**	-0.43	0.09	-0.16	-0.08
		-0.96 to -0.25	-0.90 to 0.17	-0.42 to 0.67	-0.73 to 0.41	-0.48 to 0.35
		Pr(rg>0) = 0.00	Pr(rg>0) = 0.08	Pr(rg>0) = 0.60	Pr(rg>0) = 0.31	Pr(rg>0) = 0.37
	LE	0.16	0.13	0.17	0.16	0.15
		0.12 to 0.20	0.09 to 0.16	0.13 to 0.20	0.12 to 0.19	0.12 to 0.17

For IBW, the posterior mean of the direct heritability was lower than the maternal heritability in D1, whereas the inverse was the case for D2. The 95%-HPD intervals were still large, but their lower level deviated substantially from zero, particularly for maternal heritabilities. The posterior mean of the correlation between the direct and

maternal genetic effects was negative at -0.64 and its probability to be negative was $Pr(r_g \leq 0) = 1.00$ in D1.

In contrast, D2 showed a positive genetic correlation at 0.09 between direct and maternal effects, which is expected to be positive at a probability of $Pr(r_g > 0) = 0.60$. Due to the high negative covariance between direct and maternal effects, the total heritability was low at 0.08 for D1, whereas the positive covariance between those effects for D2, in combination with similar direct and maternal variances, resulted in a moderate total heritability of 0.28.

2.3.5 Piglet level correlations

The posterior means of phenotypic correlations between SVBP and IBW were negative for both lines, with $Pr(r_p \leq 0) = 1.00$, and thus unfavourable with means of -0.53 and -0.62 for D1 and D2, respectively (Table 2.8). The posterior means of the correlation between the direct genetic effects of both traits was positive in D1 (0.62, $Pr(r_g > 0) = 0.99$), but negative in D2 (-0.42, $Pr(r_g \leq 0) = 0.97$). The posterior means of correlations between the maternal genetic effects of both traits were much less pronounced at positive level of 0.18 in D1 ($Pr(r_g > 0) = 0.82$) and at negative level of -0.17 ($Pr(r_g \leq 0) = 0.80$) in D2. The correlation between the direct genetic effect of one trait and the maternal genetic effect of the respective other trait was only different from zero for $SVBP_d$ – IBW_m in D1 with a 95%-HPD of -0.99 to -0.88. The posterior means of correlations between litter effects of SVBP and IBW were moderately positive in the range of 0.38 to 0.41 while their residual correlations were highly negative (-0.87 to -0.89 $Pr(r_g \leq 0) = 1.00$) and very consistent among the lines.

2.3.6 Adjustment for litter size

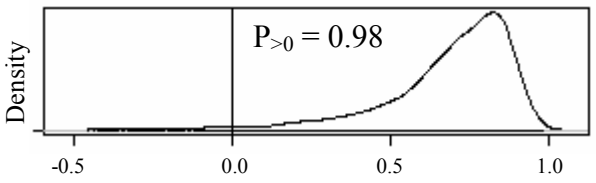
To evaluate the influence of litter size on SVBP and IBW, in a further analysis at piglet level both traits were adjusted for NBT. Adjustment for NBT led to higher posterior means of variance components for the direct and maternal genetic effect of SVBP and a slight increase of the corresponding heritabilities (Table 2.7). However, total heritabilities did not increase, mainly because the posterior means of

correlations between the direct and maternal genetic effect changed for both lines from positive to negative, which compensated the increase in direct and/or maternal genetic variance due to adjustment for litter size. Especially the combined analysis of lines D12 showed a strong negative mean posterior correlation (-0.62, $Pr(r_g>0) = 0.01$). In contrast, for IBW all posterior means of direct heritabilities decreased after adjustment for litter size. The decrease in direct genetic variance in D2 combined with the negative covariance between direct and maternal genetic effect lead to a visible decrease in total heritability in this line from 0.28 to 0.11.

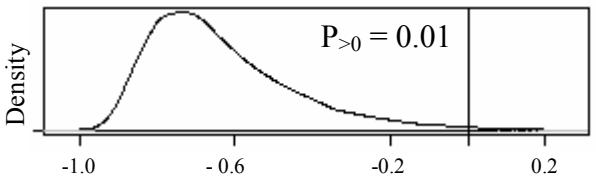
Table 2.8: Correlations at piglet level. Posterior means of genetic (r_g), litter effect (r_l), residual (r_e) and phenotypic (r_p) correlations between piglet traits in dam lines D1, D2 and D12 unadjusted (UAJ) and adjusted (AJ) for litter size (95%-HPD and posterior probability of a positive correlation Pr(r>0) as subscript; estimates of correlations with 95%-HPD excluding zero bold).

	D1		**D2**		**D12**
Trait	UAJ	AJ	UAJ	AJ	AJ
r_g $SVBP_d$ - IBW_d	**0.62**	**0.68**	**-0.42**	**-0.64**	**0.76**
	0.23 to 0.94	**0.25 to 0.98**	**-0.83 to -0.05**	**-0.91 to -0.26**	**0.57 to 0.92**
	Pr(rg>0) = 0.99	**Pr(rg>0) = 0.98**	**Pr(rg>0) = 0.03**	**Pr(rg>0) = 0.01**	**Pr(rg>0) = 1.00**
r_g $SVBP_m$ - IBW_m	0.18	0.24	-0.17	0.01	0.03
	-0.21 to 0.58	-0.16 to 0.68	-0.55 to 0.25	-0.34 to 0.38	-0.26 to 0.31
	Pr(rg>0) = 0.82	Pr(rg>0) = 0.87	Pr(rg>0) = 0.20	Pr(rg>0) = 0.50	Pr(rg>0) = 0.58
r_g $SVBP_d$ - IBW_m	**-0.94**	**-0.93**	-0.37	-0.37	-0.37
	-0.99 to -0.88	**-0.99 to -0.83**	-0.81 to 0.10	-0.86 to 0.19	-0.70 to 0.00
	Pr(rg>0) = 0.00	**Pr(rg>0) = 0.00**	Pr(rg>0) = 0.06	Pr(rg>0) = 0.12	Pr(rg>0) = 0.03
r_g $SVBP_m$ - IBW_d	-0.15	-0.23	0.25	**0.72**	-0.09
	-0.80 to 0.51	-0.82 to 0.40	-0.33 to 0.70	**0.33 to 0.99**	-0.56 to 0.44
	Pr(rg>0) = 0.33	Pr(rg>0) = 0.26	Pr(rg>0) = 0.82	**Pr(rg>0) = 1.00**	Pr(rg>0) = 0.37
r_l SVBP - IBW	**0.41**	0.39	0.38	**0.41**	**0.36**
	0.01 to 1.00	-0.07 to 1.00	0.00 to 0.80	**0.04 to 0.83**	**0.06 to 0.68**
	Pr(rl>0) = 0.95	Pr(rl>0) = 0.91	Pr(rl>0) = 0.99	**Pr(rl>0) = 0.99**	**Pr(rl>0) = 1.00**
r_e SVBP - IBW	**-0.89**	**-0.87**	**-0.87**	**-0.80**	**-0.85**
	-0.97 to -0.81	**-0.95 to -0.79**	**-0.94 to -0.80**	**-0.86 to -0.75**	**-0.90 to -0.80**
	Pr(re>0) = 0.00	**Pr(re>0) = 0.00**	**Pr(re>0) = 0.00**	**Pr(re>0) = 0.00**	**Pr(re>0) = 0.00**
r_p SVBP - IBW	**-0.62**	**-0.65**	**-0.53**	**-0.55**	**-0.59**
	-0.67 to -0.56	**-0.70 to -0.59**	**-0.60 to -0.46**	**-0.62 to -0.49**	**-0.64 to -0.55**
	Pr(rp>0) = 0.00	**Pr(rp>0) = 0.00**	**Pr(rp>0) = 0.00**	**Pr(rp>0) = 0.00**	**Pr(rp>0) = 0.00**

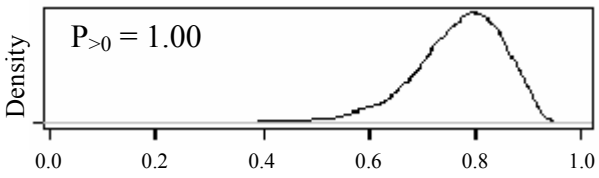
After adjustment for NBT, the posterior means of the correlations between maternal or direct effects of SVBP and direct effects of IBW mainly changed in D2 (Table 2.8). This may be due to the larger decrease of the direct heritability in this line. The marginal posterior distributions of the correlations between direct genetic effects of SVBP and IBW are shown in Figure 2.1. Even though these correlations were distinctly different from zero, estimates changed substantially whether separate or combined analyses of data from the lines were used.



(a) D1: r_g SVBP$_d$ - IBW$_d$ = 0.68 $_{0.25 \text{ to } 0.98}$



(b) D2: r_g SVBP$_d$ - IBW$_d$ = -0.64 $_{-0.91 \text{ to } -0.26}$



(c) D12: r_g SVBP$_d$ - IBW$_d$ = 0.76 $_{0.57 \text{ to } 0.92}$

Figure 2.1: Genetic correlations between survival and birth weight. Marginal posterior distribution, mean and 95%-HPD (as subscript) of the correlation between the direct genetic effects of SVBP and IBW in dam lines D1 (a), D2 (b) and D12 (c) after adjustment for NBT.

2.4 Discussion

2.4.1 Piglet survival at birth

At litter level all lines showed considerable variation for survival at birth, and the two lines with low average birth weight (D2 and S1) showed relatively high heritabilities for SVBL. This could be due to biological differences among the lines as a result of different breeding goals and genetic background (breeds), or, in the case of line S1, the low percentage of piglet survival. In the dam lines, selection for litter size and piglet survival has led to a higher survival at birth; however, across all lines the heritabilities for survival were not clearly different. For most lines in this study the heritabilities were in agreement with literature for various survival traits – such as survival at birth, survival during early pre-weaning, survival during late pre-weaning and total pre-weaning survival – which ranged from 0.05 to 0.14 (SERENIUS *et al.* 2004b; CANARIO *et al.* 2006b; ROSENDO *et al.* 2007b; SU *et al.* 2007). Breed comparison studies showed clear differences in genetic parameters among breeds (SERENIUS *et al.* 2004b; SU *et al.* 2007; SU *et al.* 2008), though most did not distinguish between sire and dam lines except for KNOL *et al.* (2002a) and IBÁÑEZ-ESCRICHE *et al.* (2009b).

At piglet level, the total genetic heritabilities of SVBP agree with those reported by SU *et al.* (2008) for Danish Landrace (0.04) and Danish Yorkshire (0.03). The much greater magnitudes of the maternal genetic effect compared to the direct genetic effect indicate that SVBP depends more on the genetic effect of the dam rather than the piglet. Studies that considered both genetic effects found heritabilities for the direct genetic effect of survival at birth ranging from 0.00 to 0.10 and for the maternal genetic effect ranging from 0.03 to 0.13 in various breeds (KNOL *et al.* 2002a; SU *et al.* 2008; IBAÑEZ-ESCRICHE *et al.* 2009b; ROEHE *et al.* 2009).

The correlation between the direct and maternal genetic effect of SVBP was favourable, but adjustment for litter size lead to a negative correlation. Correlations between the direct and maternal effect of survival (adjusted for litter size) vary

considerably in the literature, ranging from -0.56 to 0.15, but tend to be negative (KNOL *et al.* 2002a; SU *et al.* 2008; IBAÑEZ-ESCRICHE *et al.* 2009b). The influence of adjustment for litter size differed between papers (e.g. ROEHE 1999; ROEHE *et al.* 2009) and might be dependent on the population analysed. The unfavourable correlations between the direct and maternal genetic effect of survival suggest that selection for either component might compromise overall improvement if no simultaneous selection for the other genetic effect occurs.

2.4.2 Birth weight

Birth weight was measured individually on piglets that were born alive. For still born piglets the birth weight was not recorded, and therefore could not be taken into account. Consequently this might bias the results regarding weight traits, though research has shown that differences between estimates of genetic parameters for this trait with weights of still born included or excluded did not change the heritability of SVBP (ROEHE *et al.* 2009). Heritabilities for the weight traits showed considerable differences among the lines. Several studies looking at birth weight at litter level found heritabilities ranging from 0.08 to 0.43 for litter birth weight (ROEHE 1999; HERMESCH *et al.* 2000c; CHIMONYO *et al.* 2006; ROSENDO *et al.* 2007b) and heritabilities ranging from 0.15 to 0.39 for average birth weight (HERMESCH *et al.* 2000c; DAMGAARD *et al.* 2003; ROSENDO *et al.* 2007b). Heritabilities for STD were slightly higher in this study than generally found in the literature, ranging from 0.03 to 0.08 (DAMGAARD *et al.* 2003; WOLF *et al.* 2008). S1 in particular, with the lowest ALBW and STD, showed a high heritability for STD when compared to the other lines.

Estimates of heritabilities for IBW were slightly higher than those reported in the literature, where heritabilities ranged from 0.01 to 0.09 (direct) and from 0.16 to 0.22 (maternal), with genetic correlations ranging from -0.41 to 0.30 (ROEHE 1999; KNOL *et al.* 2002a; SU *et al.* 2008). The total heritability ranged from 0.08 to 0.28, a range larger than estimated by SU *et al.* (2008) (0.12 to 0.15). The unfavourable correlations between the direct and maternal genetic effect suggest that genetic

effects of the dam (i.e. uterine nutrition and capacity of the uterus) are compromising the genetic effects of the piglets (i.e. growth potential) within the litter (ROEHE 1999). This was particularly visible in D1, which had a much higher IBW than D2. Adjustment for litter size resulted in marginally more precise (smaller 95%-HPD) estimates of variances and heritabilities, which was also found by ROEHE (1999).

2.4.3 Correlation between survival and birth weight

One objective was to investigate if the additional selection for weight traits improves survival at birth rather than selection for survival *per se*. D1 had a high ALBW and STD and was the only line to show negative correlations of ALBW and STD with SVBL, similar to what is found in Swedish Yorkshire dams (DAMGAARD *et al.* 2003). The negative correlation between ALBW and SVBL suggests that, for this line, there is no necessity to increase ALBW.

Interestingly, at piglet level, both D1 and D12 showed a significantly favourable correlation between the direct genetic effects of SVBP and IBW, which suggests that as piglet birth weight increases, the chance of survival for an individual piglet also increases. This correlation is in agreement with results of SU *et al.* (2008). However, in D1 the maternal genetic effect of IBW was negatively correlated with the direct genetic effect of SVBP, which may explain the negative correlation between SVBL and ALBW estimated at litter level. The genetic effect of the dam to provide a better environment for the offspring combined with a genetic effect of piglets for an increased birth weight results in competition among piglets for limited space and resources. Piglets striving to achieve a higher birth weight do so to the detriment of their littermates, which in turn might lead to a lower survival of those littermates.

Surprisingly, at piglet level the genetic correlations in D2, which had the lowest IBW of the dam lines, were very different from the D1 line. Except for an unfavourable correlation between the maternal genetic effect of IBW and the direct genetic effect of SVBP, all correlations in this line were the opposite of D1. SU *et al.* (2008) estimated a negative correlation between the maternal genetic effects of IBW and

SVBP in Danish Landrace pigs, but both SU *et al.* (2008) and GRANDINSON *et al.* (2005) found a positive correlation in Danish and Swedish Yorkshire, respectively. This underlines that differences among breeds can be substantial, as was estimated between D1 and D2.

The significantly negative correlation between the direct genetic effects in D2 suggests that, even though IBW is already low in this line, an increase of IBW could have a negative effect on SVBP. However, the direct genetic effects of IBW showed a positive correlation to maternal genetic effects of SVBP. This may be of more importance for improvement of survival due to the higher heritability of maternal genetic effects of SVBP.

An undesirable aspect for genetic improvement of survival is the consistently negative correlation obtained between direct genetic effects of SVBP and maternal genetic effects of IBW. The correlation between the litter effects showed that permanent environmental effects of IBW within a litter have a positive correlation with SVBP. The noticeably negative residual correlation suggests that the residual variation of IBW plays a large role in the overall negative phenotypic correlation between the two traits. For both D1 and D2 the combination of favourable and unfavourable genetic correlations suggests that, for inclusion of selection for IBW, the different levels of the genetic effects (direct versus maternal) have to be given careful consideration to increase overall survival.

2.4.4 Litter size associations with survival and birth weight

The heritabilities for NBT in the present study are 0.11 to 0.16, which is similar to those reported in literature, ranging from 0.10 to 0.25 (SOUTHWOOD and KENNEDY 1990; ROEHE and KENNEDY 1995; BOUQUET *et al.* 2006). Genetic correlations between NBT and SVBL varied considerably among lines, in agreement with the large range found in literature, ranging from -0.38 to 0.29 (SERENIUS *et al.* 2004b; ROSENDO *et al.* 2007b; SU *et al.* 2007). The negative correlations between NBT and ALBW are similar to those found in literature, where litter size traits (either number

born in total or number born alive) showed clear negative correlations with ALBW, ranging from -0.30 to -0.86 (HERMESCH *et al.* 2000c; DAMGAARD *et al.* 2003; ROSENDO *et al.* 2007b).

A negative correlation, especially at phenotypic level, was expected: once the maximum uterine capacity has been reached, an increase in either NBT or ALBW will lead to a decrease in the respective other trait. An increase of NBT is associated with an increase of STD, since a larger number of piglets provides more opportunity for variation of birth weights, which was also observed in the phenotypic correlations.

The favourable genetic correlation between SVBL and NBT in D1, S1 and S2 indicates that simultaneous improvement of both litter size and survival in these lines is efficiently achievable, in agreement with KNAP (2008). The unfavourable correlation between SVBL and NBT in D2 and D3 would require dedicated selection for both traits in order to achieve simultaneous genetic improvement. Favourable correlations between SVBL and the two weight traits ALBW and STD give further genetic opportunities to achieve an increase in litter size without increase in mortality. In line D1, improvement of survival through addition of selection for STD is expected to be successful despite the low heritability of survival, due to the high correlations of both NBT and STD with SVBL. Furthermore, selection of the traits on the individual piglet level may be more successful than on litter level.

2.5 Conclusions

All lines showed considerable variation of genetic effects for survival at birth and relatively high heritabilities for this trait in lines with low average birth weight. At litter level both average birth weight and variation of birth weight within litter showed moderate heritabilities, combined with mostly favourable genetic correlations with survival. At individual piglet level, maternal heritabilities of birth weight were mostly moderate in magnitude and thus of interest for selection. Using

average birth weight at litter level might be suboptimal compared to individual birth weight at piglet level if a direct and maternal effects model is valid with negative correlations between these effects. Several highly favourable correlations of individual birth weight with survival at birth were present and selection for individual birth weight therefore seems to be a viable means of improving piglet survival at birth, provided the benefits of higher heritabilities outweigh the added costs of weighing each individual piglet.

The variation in parameters among lines indicates that the choice of the optimal inclusion of selection for birth weight might be considered for each line individually to maximise overall genetic improvement in piglet survival and growth. The current breeding goals of dam lines include litter size, which has to be taken into account when improving survival indirectly by selection for birth weight. Improvement of litter size, weight and consequently survival at litter level seems to be possible, however, there is a limit in how far both litter size and birth weight can be increased simultaneously due to their negative correlation. The differences in heritabilities among the lines also indicate that, for each line, a separate strategy to improve survival has to be considered. Selection for a weight or litter size trait may have a positive effect on survival for one line, but have no effect, or even negative effects, in another line if no stabilising selection is placed on other traits that are undesirably associated with survival.

Chapter 3 – Impact of selection for piglet survival on reproduction and production traits

3.1 Introduction

The aim of this chapter was threefold. Firstly, heritabilities of piglet survival traits were estimated, and their genetic associations with other reproduction traits – such as number of piglets born in total and number of piglets born alive – as well as production traits – such as average daily gain and backfat thickness – were estimated. Secondly, the difference in genetic parameters for reproduction and production traits in a sire line and a dam line that originated from the same breed, but differed in their breeding goal, were examined. Thirdly, by changing the base population through a combined restriction of depth of the pedigree and performance data (to recent years) it was investigated how genetic parameters and associations between traits changed within line due to the selection emphasis on different traits. A Bayesian approach was used to estimate genetic variances and covariances between traits to obtain more specific information of the precision of the estimates using the posterior distribution of the genetic parameters.

3.2 Materials and methods

3.2.1 Animals

Approximately 25 years ago, the British pig breeding organisation JSR Genetics separated their Large White breed into two different breeding lines: one line selected primarily for production traits and used as sire line; and a second line selected with greater emphasis on reproduction traits and used as dam line. For each line three data sets were available, containing information on reproductive and production performance as well as causes of piglet death. Reproductive performance data were available from April 1992 till September 2006 for the sire line (4713 litters), and from June 1990 till January 2007 for the dam line (14,836 litters). Data on production performance (i.e. information for the growing and finishing phases) were available from April 1991 till February 2007 in both the sire line (58,329 pigs) and the dam line (108,912 pigs).

In total eight different traits were considered for the analysis: two mortality traits (percentage of stillborn piglets (SB) and percentage of piglets dead from birth till weaning (DW)); three litter traits (number of piglets born in total (NBT), number of piglets born alive (NBA) and number of piglets weaned (NW)) and three production traits (average daily gain in kg/day (ADG), backfat thickness in mm (BF) and rib muscle depth in mm (MD). MD was only measured in the sire line. SB was calculated as the percentage of piglets stillborn out of the number of piglets born in total, while DW was calculated as the percentage of piglets that died from birth till weaning out of the litter size after cross-fostering (i.e. including piglets fostered on and excluding piglets fostered off). All litter size and production traits were normally distributed; survival traits showed a slightly skewed distribution but transformation was not considered to be necessary due to the large number of observations. Piglets for both lines entered the performance test on average at an age of 95 days. Piglets in the sire line weighed on average 44 kg at the start of the test and 90 kg at the end of the test, and spent on average 54 days on test; piglets in the dam line weighed on average 43 kg and 91 kg, respectively, and were on average 55 days on test. The performance test was between 40 kg and 91 kg so that ADG was adjusted for small differences from both of these targets weights and BF and MD were adjusted for an end of test weight of 91 kg. The two mortality and three litter size traits will be referred to as reproduction traits and the three growing finishing traits referred to as production traits.

Furthermore, the data sets contained information on several systematic effects, namely batch, service type, parity, gestation length and weaning period. Batches based on farrowing-unit, year and season were fitted in the model for reproduction performance. Observations for the sire line came from four different farrowing units and for the dam line from 11 farrowing units, whereby three units were present in both lines. Management practices were standardised across the organisation and hence did not differ between units. The seasonal effect was determined by splitting a 12 month period in two seasons, April-September and October-March. Batches based on sex, production unit, year and season, which was the year divided into quarterly

seasons, were fitted in the model for production performance. Animals for the production data came from three different production units for the sire line and five for the dam line, whereby one unit was present in both lines. Service type was either natural service, on farm AI, or AI from an AI station. In the sire line parities one to five were considered as separate classes and all sows with six or more parities were grouped together. In the dam line parities one to seven were considered as separate classes and all sows with eight and more parities were grouped together. Gestation length in days was grouped as $\leq 111, 112, ..., 118, \geq 119$ for both lines. Weaning period in days was grouped as $\leq 16, 17, ..., 35, \geq 36$ for the sire line and $\leq 11, 12, ..., 36, \geq 37$ for the dam line. Cross-fostering was applied two days after birth and occurred in 42% of the litters in the sire line and 37% of the litters in the dam line, with on average three piglets per litter cross-fostered in these litters. Whenever possible litters were cross-fostered up or down into groups of twelve, or the closest possible arrangement, and aimed to minimise the number of piglets moved and to mix animals of similar size. Cross-fostering practises were consistent across farrowing units and piglets were only cross-fostered onto sows of the same genetic line, so no cross-fostering occurred between the animals in the two lines. Information regarding cross-fostering in the data set was restricted to the number of piglets fostered on or off per sow, without information of their biological mother or nurse sow, respectively, and could therefore not be accounted for in reproduction traits measured at weaning.

Connectedness of the data sets was high; 70% (500 out of 711) of the sires in the sire line had offspring in both the reproduction and production data set, accounting for 98% of the litters and 87% of the animals with production records; 67% (718 out of 1065) of the sires in the dam line had offspring in both data sets, accounting for 92% of the litters and 93% of the animals with production records. The 4713 litters in the sire line were from 2928 sows, with 60% of the sows having only one litter in the reproduction data set and the remaining 40% of the sows up to seven litters (average 1.6 parities per sow). Production data were available for 1924 of these 2928 sows (66%). The 14,836 litters in the dam line were from 7724 sows and 49% of the sows

had only one litter in the reproduction data set, whilst 51% of the sows had up to eight litters (average 1.9 parities per sow). Production performance data were available for 5504 of these 7724 sows (71%).

3.3.2 Statistical analysis

Fixed effects were tested for significance using the procedure MIXED (SAS 2002). Based on this preliminary analysis, the following fixed effects were included in the models for corresponding traits: batch effect for the production traits ADG, BF and MD; batch, service type, parity and gestation length for the reproduction traits NBT, NBA and SB; batch, service type, parity, gestation length and weaning period for NW and DW. The following models were analysed using a Bayesian approach using Gibbs sampling and MCMC methods to estimate genetic parameters for the reproduction traits

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{Wc} + \boldsymbol{e}, \quad (1)$$

and for the production traits

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{e}, \quad (2)$$

where $\boldsymbol{y}$ is the vector of observations of the traits, $\boldsymbol{b}$ the vector of fixed effects (including effects described earlier), $\boldsymbol{a}$ the vector of additive genetic effects, $\boldsymbol{c}$ the vector of the permanent environmental effects of the sow and $\boldsymbol{e}$ the vector of residuals. $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are incidence matrices relating the vectors $\boldsymbol{b}$, $\boldsymbol{a}$, and $\boldsymbol{c}$ with $\boldsymbol{y}$. For the multiple trait analysis models (1) and (2) were combined. The assumed (co)variance structure for reproduction data was

$$V\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{c}\\ \boldsymbol{e}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}\otimes\boldsymbol{G} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\otimes\boldsymbol{C} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\otimes\boldsymbol{R}\end{bmatrix},$$

where $\boldsymbol{A}$ and $\boldsymbol{I}$ are the additive genetic relationship matrix and identity matrix, respectively. $\boldsymbol{G}$, $\boldsymbol{C}$ and $\boldsymbol{R}$ represent the variance and covariance matrices of direct additive genetic effects, permanent environmental effects of the sow and residual environmental effects, respectively. For production traits the permanent environmental effects and its variances need not be considered because those traits were measured only once.

Pedigree files were checked with Relax2 (STRANDÉN and VUORI 2006) for cycles, missing animals and consistency. No errors were found and after checking, the pedigree files were matched to the animals in the data set to eliminate superfluous animals in the pedigree. Pedigrees for animals in the sire line were traced back as far as 1987, while pedigrees for animals in the dam line were traced back to 1985 and no overlap between animals in the two pedigree files occurred. No limit was set for the number of generations included in the pedigree files, so depending on the birth year of the animal up to sixteen generations were available. The two pedigree files contained 60,021 and 112,205 animals for the sire and dam line, respectively. Records in the sire line included 602 sires and 3304 dams with offspring; records in the dam line included 800 sires and 6236 dams with offspring.

Data sets were analysed based on a Bayesian approach using Gibbs sampling and MCMC methods with flat priors using the programme GIBBS2F90 (MISZTAL *et al.* 2002). Due to computational limitation the traits were first genetically analysed in two groups per line, one group containing the two mortality traits and three litter traits, the other group containing the production traits. After some exploratory analyses, chains of 500,000 to 800,000 samples were used, depending on the (combination of) traits, with a burn-in of 50,000 to 250,000 and a lag of 50. Thus marginal posterior distributions were estimated with a minimum of 5000 samples each. Convergence was tested using the Geweke criterion (GEWEKE 1992), Raftery and Lewis criterion (RAFTERY and LEWIS 1992) and visual assessment of the drawn marginal posterior distributions.

Cross-fostering from birth till weaning could not be accounted for and models containing NW and DW often did not converge, therefore these traits were excluded from further analyses. Current cross-fostering practices have a high influence on genetic parameters for traits measured at weaning and not applying cross-fostering could be of advantage to estimate reliable parameters for these traits. Genetic parameters of traits were then estimated in one single multiple trait analysis, containing all six traits in the sire line and all five traits in the dam line. Single trait analyses showed that this multiple trait analysis did not inflate the phenotypic variances (unpublished results). In order to examine the change of genetic parameter within line, in a further analysis the data set was restricted to only include records and litters of animals born in the year 2002 or later, which will be referred to as the 'restricted data set' as compared to the 'full data set'. Additionally, the pedigree was restricted to animals born in the year 2000 and later, to change the base population from the year 1985 to the year 2000. Genetic analysis of the restricted data sets was performed for the dam line only due to the low number of records in the sire line.

3.3 Results

3.3.1 Descriptive results

The mean percentage of stillborn piglets and piglets that died from birth till weaning was 8.0% and 18.2%, respectively, in the sire line, and 7.3% and 17.6%, respectively, in the dam line (Tables 3.1 and 3.2). The differences of these traits between the two lines were significant at $P < 0.001$ for SB and $P < 0.05$ for DW. In the restricted data set, SB was still 8.0% in the sire line, but slightly increased in the dam line at 7.7%. DW increased in the sire line by approximately 1%-point to 19.1%, while it decreased in the dam line by almost 3%-point to 14.9% ($P < 0.001$). Litter sizes at different stages were also significantly different between the two lines. The dam line, selected mainly for reproductive performance, had on average 12.0 piglets born in total and 11.1 piglets born alive, and approximately 1.5 more piglets per litter born than the sire line ($P < 0.001$ for NBT and NBA). The difference between these two lines in number of piglets weaned was slightly less, with 0.8 more

piglets per litter weaned in the dam line than in the sire line ($P < 0.001$). In the restricted data set means for NBT, NBA and NW increased compared to the full data set, but showed a higher increase in the dam line compared to the sire line. Consequently, the difference between the lines increased to approximately 1.7 piglets per litter for all three traits ($P < 0.001$ for NBT, NBA and NW).

Differences in ADG were small, 0.86 kg per day in the sire line compared to 0.87 kg per day in the dam line, but still significantly different between the lines ($P < 0.001$). In the restricted data set the ADG in the sire line increased to 0.92 kg/day, while the ADG in the dam line stayed at almost the same magnitude. Selection pressure on productive performance in the sire line has led to a significantly lower value for BF in this line (8.8 mm) compared to the dam line (10.5 mm) ($P < 0.001$), but this difference was less notable, though still highly significant ($P < 0.001$), in the restricted data set (9.2 mm in the sire line versus 9.9 mm in the dam line). This reflects the changing emphasis in the selection index, with a reduced emphasis on this trait once the average backfat thickness is below 10 mm in the UK.

Table 3.1: Summary statistics for the full data set. Descriptive statistics of the traits using all information and the significance of the difference (s.d.) between the lines (N = number of records, s.e. as subscript of the mean).

Trait	Sire line			Dam line			SD
	N	Mean	s.d.	N	Mean	s.d.	
SB	4713	8.04 $_{0.162}$	12.299	14,836	7.33 $_{0.091}$	10.694	***
DW	4713	18.18 $_{0.246}$	17.351	14,836	17.58 $_{0.139}$	16.740	*
NBT	4713	10.48 $_{0.049}$	3.163	14,836	12.01 $_{0.027}$	3.405	***
NBA	4713	9.66 $_{0.047}$	3.129	14,836	11.10 $_{0.026}$	3.214	***
NW	4713	8.20 $_{0.034}$	2.209	14,836	8.99 $_{0.019}$	2.336	***
ADG	58,329	0.86 $_{0.001}$	0.150	108,912	0.87 $_{0.000}$	0.131	***
BF	58,329	8.82 $_{0.008}$	1.979	108,912	10.47 $_{0.006}$	2.013	***
MD	58,329	43.58 $_{0.033}$	8.050	-	-		-

Table 3.2: Summary statistics for the reduced data set. Descriptive statistics of the traits using restricted to animals born in the year 2002 or later all and the significance of the difference (s.d.) between the lines (N = number of records, s.e. as subscript of the mean).

Trait	Sire line			Dam line			SD
	N	Mean	s.d.	N	Mean	s.d.	
SB	748	7.97 $_{0.392}$	10.796	2475	7.65 $_{0.216}$	10.714	
DW	748	19.05 $_{0.480}$	16.338	2475	14.92 $_{0.264}$	12.002	***
NBT	748	11.31 $_{0.125}$	3.053	2475	13.08 $_{0.068}$	3.506	***
NBA	748	10.42 $_{0.117}$	3.021	2475	12.03 $_{0.064}$	3.258	***
NW	748	8.27 $_{0.065}$	2.061	2475	9.99 $_{0.036}$	1.698	***
ADG	9231	0.92 $_{0.001}$	0.144	23,877	0.86 $_{0.001}$	0.116	***
BF	9231	9.23 $_{0.019}$	1.899	23,877	9.92 $_{0.012}$	1.848	***
MD	9231	55.81 $_{0.099}$	9.489	-	-		-

3.3.2 Separate analysis of production and reproduction traits

In the first genetic analysis, traits were analysed in two groups per line, one group containing the two mortality and three litter size traits, the other group containing the production traits. Tables 3.3 and 3.4 present the additive (σ^2_d), permanent environmental (σ^2_{pe}) and total phenotypic (σ^2_p) variances with standard errors for all eight traits in the sire and dam line. The reproductive traits (SB, DW, NBT, NBA and NW) had repeated measurements of traits as opposed to the production traits (ADG, BF and MD), which were measured only once per animal. Therefore, a permanent environmental variance based on the sow was included for the five reproduction traits, while none was included for the production traits. MD was only measured in the sire line.

Phenotypic variances for the production trait ADG were similar in both the sire and dam line (less than 0.01% difference), while the phenotypic variance for BF was 16% higher in the sire line than in the dam line. The phenotypic variances for NBT and NBA were 11% and 5% lower in the sire line, respectively, and for NW 15%

higher in the sire line. For SB and DW there were large differences. Total phenotypic variance for SB was 33% higher in the sire line than in the dam line, with a substantially higher residual variance in the sire line but additive genetic and permanent environmental variances that were twice as high in the dam line as in the sire line. Phenotypic variance for DW was 39% higher in the sire line than in the dam line. The additive genetic and permanent environment variances of DW in the dam line were very small compared to those in the sire line, which may be influenced by cross-fostering of piglets.

Table 3.3: Variance components and heritabilities in the sire line. Posterior means of additive genetic (σ^2_d), permanent (σ^2_{pe}) and phenotypic variance (σ^2_p), heritability (h^2) and phenotypic proportion of the permanent environmental effect (PE) for all traits in the sire line (95%-HPD as subscript).

Trait	σ^2_d	σ^2_{pe} [a]	σ^2_p	h^2	PE [a]
SB	4.87	2.13	147.5	0.03	0.01
	1.58 to 7.99	0.46 to 3.61	141.4 to 153.3	0.01 to 0.05	0.00 to 0.02
DW	22.49	55.45	279.5	0.08	0.20
	11.21 to 35.85	39.27 to 72.59	267.0 to 292.6	0.04 to 0.13	0.14 to 0.26
NBT	1.44	1.03	9.13	0.16	0.11
	0.90 to 1.99	0.53 to 1.49	8.71 to 9.54	0.10 to 0.22	0.06 to 0.16
NBA	1.27	0.73	8.92	0.14	0.08
	0.80 to 1.78	0.38 to 1.11	8.53 to 9.33	0.09 to 0.19	0.04 to 0.12
NW	0.33	0.75	4.48	0.07	0.17
	0.17 to 0.51	0.52 to 0.99	4.27 to 4.67	0.04 to 0.11	0.12 to 0.22
ADG	0.004	-	0.014	0.31	-
	0.004 to 0.005		0.014 to 0.015	0.29 to 0.34	
BF	1.90	-	3.65	0.52	-
	1.78 to 2.02		3.58 to 3.72	0.50 to 0.55	
MD	11.57	-	26.77	0.43	-
	10.80 to 12.35		26.30 to 27.22	0.41 to 0.45	

[a] Permanent environmental effect only included for reproduction traits

Heritability estimates for reproduction traits were overall low, both in the sire and dam line (Tables 3.3 and 3.4). In general these heritabilities were slightly higher in the sire line than in the dam line (0.08 vs. 0.00 for DW, 0.16 vs. 0.12 for NBT, 0.14 vs. 0.10 for NBA and 0.07 vs. 0.01 for NW). In contrast, the heritability for SB in the dam line (0.07) was higher than the heritability in the sire line (0.03).

Heritability estimates for production traits were substantially higher than those of reproduction traits. These heritabilities were moderate to high, ranging from 0.30 for ADG in the dam line to 0.52 for BF in the sire line. Heritabilities for ADG were similar for both lines (0.31 vs. 0.30) but for BF the heritability in the sire line (0.52) was much higher than in the dam line (0.42).

Table 3.4: Variance components and heritabilities in the dam line. Posterior means of additive genetic (σ^2_d), permanent (σ^2_{pe}) and phenotypic variance (σ^2_p), heritability (h^2) and phenotypic proportion of the permanent environmental effect (PE) for all traits in the dam line (95%-HPD as subscript).

Trait	σ^2_d	σ^2_{pe} [a]	σ^2_p	h^2	PE [a]
SB	8.26	4.44	111.1	0.07	0.04
	5.50 to 11.08	1.82 to 7.24	108.4 to 113.9	0.05 to 0.10	0.02 to 0.06
DW	0.38	3.57	201.5	0.00	0.02
	0.20 to 0.58	0.25 to 7.13	197.1 to 206.1	0.00 to 0.00	0.00 to 0.04
NBT	1.20	1.10	10.25	0.12	0.11
	0.90 to 1.50	0.83 to 1.40	9.98 to 10.50	0.09 to 0.14	0.08 to 0.14
NBA	0.97	0.97	9.43	0.10	0.10
	0.71 to 1.26	0.72 to 1.23	9.21 to 9.67	0.08 to 0.13	0.07 to 0.13
NW	0.04	0.08	3.91	0.01	0.02
	0.01 to 0.06	0.02 to 0.14	3.82 to 4.00	0.00 to 0.02	0.00 to 0.03
ADG	0.004	-	0.014	0.30	-
	0.004 to 0.005		0.014 to 0.015	0.28 to 0.31	
BF	1.33	-	3.14	0.42	-
	1.26 to 1.40		3.10 to 3.18	0.41 to 0.44	

[a] Permanent environmental effect only included for reproduction traits

3.3.3 Combined analysis using all pedigree information

Tables 3.5 and 3.6 present the results of the genetic and phenotypic correlations of the combined analysis of production and reproduction traits for the sire and dam line excluding the traits NW and DW because of influence due to cross-fostering. Heritabilities based on this analysis, which included all production and reproduction traits, were slightly lower than those estimated in separate analyses of reproduction (1) and production (2) traits. Genetic and phenotypic correlations between the two litter size traits were high, ranging from 0.92 to 0.97, while genetic and phenotypic

correlations between SB and the two litter size traits were by and large not significantly different from zero.

Table 3.5: Correlations in the sire line using the full data set. Heritabilities (diagonal, bold), genetic correlations (above diagonal) and phenotypic correlations (below diagonal) for traits of the sire line (95%-HPD as subscript) based on a base population of animals born in 1987 and observations of animals born in 1991 or later.

Trait	SB	NBT	NBA	ADG	BF	MD
SB	**0.03**	0.06	-0.17	-0.01	-0.46	0.13
	0.01 to 0.05	-0.40 to 0.53	-0.60 to 0.30	-0.28 to 0.26	-0.74 to -0.20	-0.12 to 0.38
NBT	-0.03	**0.10**	0.97	-0.03	0.10	-0.19
	-0.06 to 0.00	**0.05 to 0.15**	0.94 to 0.99	-0.19 to 0.13	-0.05 to 0.27	-0.36 to -0.03
NBA	-0.38	0.92	**0.09**	-0.04	0.18	-0.23
	-0.40 to -0.35	0.92 to 0.93	**0.04 to 0.14**	-0.21 to 0.13	0.01 to 0.35	-0.41 to -0.04
ADG	0.00	0.00	-0.01	**0.29**	0.28	-0.14
	-0.02 to 0.02	-0.03 to 0.02	-0.03 to 0.02	**0.27 to 0.31**	0.23 to 0.33	-0.20 to -0.08
BF	-0.05	0.02	0.04	0.15	**0.50**	-0.30
	-0.09 to -0.02	-0.01 to 0.06	0.00 to 0.07	0.14 to 0.16	**0.47 to 0.52**	-0.35 to -0.26
MD	0.00	-0.04	-0.04	-0.06	-0.24	**0.41**
	-0.01 to 0.04	-0.07 to -0.01	-0.07 to -0.01	-0.07 to -0.05	-0.25 to -0.23	**0.39 to 0.43**

Table 3.6: Correlations in the dam line using the full data set. Heritabilities (diagonal, bold), genetic correlations (above diagonal) and phenotypic correlations (below diagonal) for reproduction and production traits of the dam line (95%-HPD as subscript) based on a base population of animals born in 1985 and observations of animals born in 1990 or later.

Trait	SB	NBT	NBA	ADG	BF
SB	**0.06**	0.21	-0.13	0.19	-0.08
	0.05 to 0.08	-0.01 to 0.43	-0.36 to 0.08	0.08 to 0.32	-0.20 to 0.04
NBT	0.07	**0.11**	0.94	0.05	-0.07
	0.06 to 0.09	**0.08 to 0.13**	0.91 to 0.96	-0.04 to 0.16	-0.16 to 0.02
NBA	-0.30	0.92	**0.09**	-0.01	-0.07
	-0.31 to -0.28	0.91 to 0.92	**0.07 to 0.12**	-0.12 to 0.09	-0.16 to 0.04
ADG	0.03	0.01	0.00	**0.29**	0.07
	0.01 to 0.04	-0.01 to 0.03	-0.02 to 0.01	**0.27 to 0.30**	0.03 to 0.11
BF	-0.01	-0.02	-0.01	0.03	**0.42**
	-0.03 to 0.01	-0.03 to 0.00	-0.03 to 0.01	0.02 to 0.04	**0.40 to 0.43**

Both the genetic and phenotypic correlations between the production traits ADG and BF were higher in the sire line than in the dam line (0.28 vs. 0.07 and 0.15 vs. 0.03,

respectively). All genetic and phenotypic correlations among production traits were unfavourable, except the correlations between BF and MD.

Genetic correlations of reproduction traits with production traits in the sire line were generally not significantly different from zero, except for the genetic correlations of BF with SB and NBA and of MD with NBT and NBA. In the dam line only the unfavourable genetic correlation of SB with ADG was significantly different from zero. The genetic correlation between SB and BF was negative in both lines, though much more pronounced in the sire line (-0.46) than in the dam line (-0.08).

3.3.4 Combined analysis using restricted pedigree information

In order to identify the change of genetic parameters within line, the base population of the dam line was changed to pigs born in the year 2000, which resulted in some different genetic parameters compared to the full data set (Table 3.7). Heritabilities for reproduction traits were again low; 0.05, 0.07 and 0.07 for SB, NBT and NBA, respectively. These heritabilities were slightly lower than the heritabilities in the full data set. Heritabilities for production traits were similar to those in the full data set.

Table 3.7: Correlations in the dam line using the restricted data set. Heritabilities (diagonal, bold), genetic correlations (above diagonal) and phenotypic correlations (below diagonal) for reproduction and production traits of the dam line (95%-HPD between parentheses) based on data restricted to a base population of animals born in 2000 and observations of animals born in the year 2002 or later.

Trait	SB	NBT	NBA	ADG	BF
SB	**0.05**	0.29	-0.03	0.30	0.14
	0.01 to 0.09	-0.30 to 0.87	-0.74 to 0.64	-0.01 to 0.64	-0.21 to 0.52
NBT	0.07	**0.07**	0.93	0.21	-0.01
	0.03 to 0.11	**0.03 to 0.12**	0.84 to 1.00	-0.05 to 0.50	-0.31 to 0.28
NBA	-0.30	0.91	**0.07**	0.09	0.03
	-0.33 to -0.26	0.91 to 0.92	**0.02 to 0.12**	-0.18 to 0.39	-0.27 to 0.32
ADG	0.03	0.03	0.01	**0.29**	0.01
	0.00 to 0.07	-0.01 to 0.07	-0.03 to 0.05	**0.26 to 0.33**	-0.07 to 0.09
BF	0.02	0.00	0.00	0.00	**0.45**
	-0.03 to 0.07	-0.06 to 0.05	-0.05 to 0.05	-0.02 to 0.02	**0.42 to 0.49**

Correlations of NBA with NBT were, as before in the full data set, highly positive, but correlations of SB with NBT and NBA and correlations between the two production traits were not significant. Genetic correlations of reproduction traits and production traits showed more desirable genetic associations, though none of them were significant. Phenotypic correlations were generally in the same direction as genetic correlations but of lower magnitude. Restricted data of the sire line were too small to result in reliable estimates and are therefore not presented.

3.4 Discussion

All analyses in this chapter were carried out using a Bayesian approach in order to obtain information about the precision of the estimation of the genetic parameter as given as Bayesian confidence intervals. Depending on the trait (or combination of traits for correlations), varying chain lengths were used, with longer chains for correlations of reproduction traits with production traits due to the difference in records (sow/litter versus individual animal information) which increased the time needed for convergence. Due to the large number of traits, not all correlations could be estimated in one single multiple trait analysis per line, and therefore had to be estimated separately for production and reproduction traits. Due to computational limits, few studies in the past have used a Bayesian approach to obtain genetic parameters for data sets of this size and even now analysis of data sets of this large size took several weeks to be completed.

The present chapter is unique in the fact that it is based on data from a sire and dam line originating from the same Large White population, divergently selected commencing 25 years ago. Few studies have differentiated between a sire and dam line, and none of them originating from the same breed (KNOL *et al.* 2002a). Studies analysing and comparing different breeds showed that there are clear differences in genetic parameters between breeds (SEE *et al.* 1993; ROEHE and KENNEDY 1995; SU *et al.* 2008). Only a few studies (FERRAZ and JOHNSON 1993; HERMESCH *et al.* 2000c; SERENIUS *et al.* 2004b) have used data from Large White pigs in their studies

to compare different breeds, and none of these studies differentiated between sire and dam line within the breed. The restriction of the data set of the dam line showed the effect of the change of parameters given a more recent base population. The change of parameters is expected to be due to change in depth of pedigree and due to use of only recent performance data.

3.4.1 Reproduction traits

Heritabilities for NBT and NBA were 0.10 and 0.09 in the sire line respectively and 0.11 and 0.09 in the dam line. Heritabilities for NBT in literature have varied considerably, ranging from 0.05 to 0.24 (SERENIUS *et al.* 2003; SU *et al.* 2007; RYDHMER *et al.* 2008), with most values around 0.10. Moreover, heritabilities of 0.09 for NBA are in accordance with heritabilities for this trait in literature, which range from 0.05 to 0.16 (ROSENDO *et al.* 2007b; SU *et al.* 2007; FERNANDEZ *et al.* 2008).

In this chapter piglet survival was defined as percentage of stillborn piglets. As opposed to NBT and NBA, a clear comparison with other studies is more difficult, since piglet mortality or its inverse piglet survival is not always defined in the same way as in this chapter. Also, survival at birth and survival at various stages pre-weaning are generally considered to be different traits, but not always treated as such. Heritabilities reported in the literature for various survival traits – such as survival at birth, survival during early pre-weaning, survival during late pre-weaning and total pre-weaning survival – range from 0.01 to 0.13 (ROSENDO *et al.* 2007b; SU *et al.* 2007; ROEHE *et al.* 2009). Only KNOL *et al.* (2002a) have reported differences in heritabilities between a sire and dam line and estimated heritabilities for survival at birth of 0.00 to 0.04 and 0.01 to 0.05 in the dam and sire line respectively, and for pre-weaning survival of 0.04 and 0.01 respectively. In this chapter, heritability for percentage of stillborn piglets was lower in the sire line ($0.03_{\ 0.01\text{-}0.05}$) than in the dam line ($0.06_{\ 0.05\text{-}0.08}$; see Figure 3.1).

In several studies survival is treated as a character of the piglet and survival traits are analysed at the piglet level, distinguishing between a direct and a maternal genetic effect. Generally estimates for the maternal genetic effects in these studies are higher than the direct genetic effect (GRANDINSON *et al.* 2005; RYDHMER *et al.* 2008; SU *et al.* 2008). In this chapter, survival as percentage of stillborn piglets was analysed at sow level, since individual piglet information was not available. These estimated heritabilities were in the same range as those from other studies that analysed survival at the sow level, but using numbers of stillborn piglets as the trait (0.02 to 0.12) (HANENBERG *et al.* 2001; SERENIUS *et al.* 2004a; SERENIUS *et al.* 2004b).

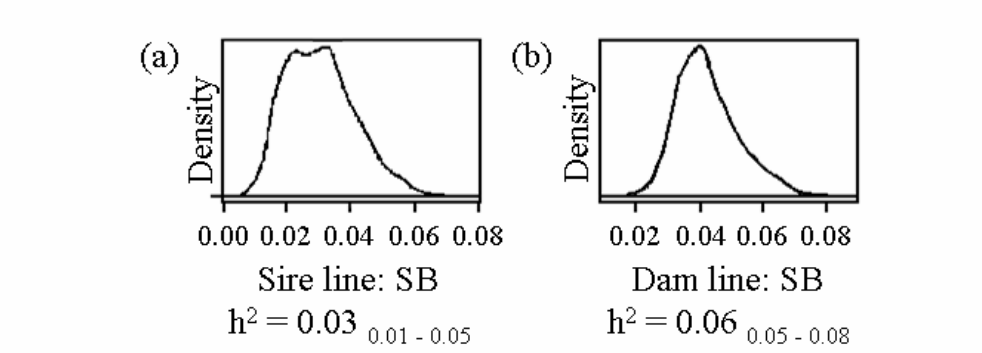


Figure 3.1: Heritability for survival. Marginal posterior distribution, mean (h2), and highest posterior density interval (as subscript) of the heritability for SB in the sire line (a) and dam line (b)

Genetic correlations between the two litter traits NBT and NBA were 0.97 and 0.94 in the sire and dam line, respectively. These correlations are at the upper end of previously reported correlations which ranged from 0.87 to 0.97 (ROEHE and KENNEDY 1995; BOUQUET *et al.* 2006; CHIMONYO *et al.* 2006). Phenotypic correlations between NBT and NBA were 0.92 in both lines, slightly higher than those in literature, ranging from 0.87 to 0.88 (BOUQUET *et al.* 2006; CHIMONYO *et al.* 2006). The genetic correlation of SB with NBT in the dam line was unfavourable at 0.21, while the same correlation in the sire line was not significantly different from zero. SU *et al.* (2007) found higher genetic correlations, ranging from -0.28 and -0.38 for the genetic correlation between percentage survival at birth and total number born

in Landrace and Yorkshire, respectively. SERENIUS *et al.* (2004b) found an unfavourable genetic correlation of 0.29 between percentage stillborn and total number born in Landrace pigs, but no significant correlation in Large White pigs, similar to the results that were obtained in the sire line. In contrast, ROSENDO *et al.* (2007b) found a favourable correlation of -0.37 between percentage stillborn and total number born in Large White pigs. The genetic correlation of SB with NBA was more similar between the two lines than the correlations for SB-NBT, with -0.17 and -0.13 for the sire and dam line respectively. Estimates of the correlation between SB and NBA reported in the literature are generally favourable, varying from -0.15 to -0.27 for the correlation between percentage stillborn and number born alive (SERENIUS *et al.* 2004b) to 0.41 to 0.61 for the correlation between percentage survival at birth and number born alive (SU *et al.* 2007).

This chapter shows that selection pressure on litter size in the dam line may have resulted in the higher undesirable correlation between NBT and SB in the dam line (Figure 3.2 (a)) as compared to the sire line. Heritabilities for reproduction traits in the restricted data set were slightly lower than in the full data set. The change of genetic correlations due to selection was small.

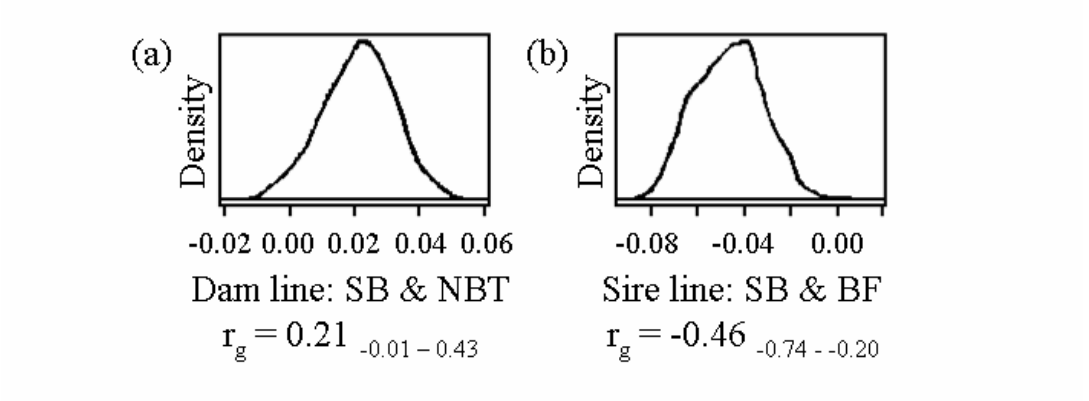


Figure 3.2: Genetic correlations between survival and litter size or backfat. Marginal posterior distributions, means (rg), and highest posterior density intervals (as subscript) of correlations between SB and NBT in the dam line (a) and between SB and BF in the sire line (b)

3.4.2 Production traits

Heritabilities for production traits were estimated at 0.29, 0.50 and 0.41 for ADG, BF and MD, respectively, in the sire line and 0.29 and 0.42 for ADG and BF, respectively, in the dam line. Heritabilities for ADG and BF were similar to values in literature which ranged from 0.23 to 0.40 for ADG (FERRAZ and JOHNSON 1993; SERENIUS and STALDER 2004; ROSENDO *et al.* 2007a) and from 0.30 to 0.51 for BF (KNOL 2001; SERENIUS and STALDER 2004; ZUMBACH *et al.* 2007). The slightly higher heritability for BF in the sire line compared to the dam line was unexpected because breeding in the sire line primarily focused on reduction of backfat thickness, while this was of lesser emphasis in the dam line. The heritability for MD was slightly higher than previously reported heritabilities which ranged from 0.12 to 0.31 (HERMESCH *et al.* 2000a; ZUMBACH *et al.* 2007).

Genetic and phenotypic correlations among the production traits were approaching zero in the dam line. In the sire line, genetic correlations for ADG-BF (0.28) and ADG-MD (-0.14) were slightly unfavourable while the genetic correlation of BF with MD was slightly favourable (-0.30). SERENIUS *et al.* (2004b) found correlations between ADG and BF of 0.32 and 0.39 in Landrace and Large White pigs, respectively, and HERMESCH *et al.* (2000b) found a slightly lower but still favourable correlation between BF and MD of -0.16 in Landrace and Large White boars. Contrary to this chapter, HERMESCH *et al.* (2000b) based their ADG on the age of the pig, where in this chapter it was based on weight of the pig. They distinguished between ADG from 3 to 18 weeks of age and ADG from 18 to 22 weeks of age and found desirable correlations between ADG from 3 to 18 weeks and both BF and MD, but undesirable correlations between ADG from 18 to 22 weeks and BF and MD. In the restricted data set of the dam line, in which the selection pressure on BF was much lower than in the sire line, the phenotypic and genetic correlations between ADG and BF showed a slight decrease.

3.4.3 Correlations of survival with reproduction and production traits

Genetic and phenotypic correlations between reproduction traits and production traits were generally low. Selection pressure on BF in the sire line may have resulted in the moderately negative correlation between SB and BF (Figure 3.2 (b)), which was much lower (but still negative) in the dam line, and the unfavourable correlations for BF-NBT and BF-NBA, which were favourable in the dam line. KNOL (2001) found unfavourable correlations of BF with both pre-weaning survival and piglet survival (defined as accumulated farrowing survival and pre-weaning survival) of 0.52 and 0.18, respectively, in a commercial sire line, similar to the correlation that were estimated in the sire line.

Correlations of ADG with reproduction traits were more pronounced in the dam line than in the sire line, with a low undesirable correlation of ADG with SB in the dam line, which was slightly favourable in the sire line, but correlations of ADG with the two litter size traits were low in both lines. Correlations between ADG and litter size traits in the literature are generally more distinctly negative, with correlations of ADG with NBA up to -0.42, depending on sow parity (HERMESCH *et al.* 2000c). Additionally, SERENIUS *et al.* (2004a) found a favourable correlation between the number of stillborn piglets and ADG, while KNOL (2001) found a favourable correlation between piglet survival (farrowing and pre-weaning survival combined) and ADG, but an unfavourable correlation of pre-weaning survival with ADG. Selection pressure on litter size in the dam line may have resulted in these more pronounced correlations with ADG in this analysis as compared to the low correlations in the sire line.

In the restricted data set, genetic correlations for ADG with litter size traits were slightly to moderately desirable, while the genetic correlation with the mortality trait SB was undesirable. For BF surprisingly most correlations were favourable except BF-NBA.

Comparison of the full data set and the restricted data set showed how selection pressure on different traits has lead to a change in heritabilities and correlations in the dam line. However, low genetic correlations between traits showed that selection pressure on either production traits or reproduction traits still leaves room for improvement of the other.

3.5 Conclusions

Genetic improvement of piglet survivability without significant reductions in performance traits is possible. Heritabilities for survivability and reproduction traits were low, but genetic variation was substantial in these traits and extensive pedigree information can be used to improve the accuracy of breeding values so that genetic improvement is expected to be efficient. Selection for reproduction traits such as number born alive will lead to improvement in survival at birth. Genetic correlations between reproduction and production traits were often undesirable in the sire line, except for a weak favourable correlation of SB with ADG. In the dam line most correlations were favourable, though some slightly unfavourable correlations were also present. The unfavourably correlated responses of SB and NBT (dam line) and SB and BF (sire line) indicate the importance of selecting for NBA in the dam line and suggests a reduced emphasis of selection for backfat thickness in combination with stabilising selection for a trait such as piglet survival in the sire line. However, in particular in the dam line, undesirable correlations between these traits were relatively low, so that simultaneous improvement of performance traits as well as piglet survival at birth can be achieved.

Chapter 4 – Performance of genomic selection

4.1 Introduction

The aim of this chapter was to assess the efficiency of genomic selection with various percentages of SNP markers considered to have an effect in a real data set, in this case mouse data. Due to the large range of phenotypes available, these data allow for an evaluation of the influence of a) the heritability of the trait, b) the QTL-distribution of the trait and c) the type of the trait ('classical' traits that are easily measurable versus behavioural traits) on the efficiency of genomic selection when different numbers of SNP markers are used. These characteristics are studied using different selection criteria (selection within family versus selection between families) and models (polygenic effects, genomic effects and a combination of both).

4.2 Materials and methods

4.2.1 Animals and SNPs

Data on 2188 geno- and phenotyped mice provided by the Wellcome Trust Centre for Human Genetics were used to analyse the efficiency of genomic selection in seven different traits. The data are freely available at http://gscan.well.ox.ac.uk/ and the population has been described and analysed comprehensively for other objectives than those in the present chapter in various papers including SOLBERG *et al.* (2006) and VALDAR *et al.* (2006b). Therefore, only the aspects important for the present analysis will be highlighted here. The animals were obtained from crossing eight purebred mice strains, followed by 50 generations of pseudo-random mating. Data comprised 175 full-sib families, collected over a period of three years, with a pedigree that consisted of parents and grandparents (2890 animals in total). The extent of linkage disequilibrium (LD) between pairs of markers was low with an $r^2 < 0.5$ within 2 Mb and < 0.2 within 8 Mb (VALDAR *et al.* 2006a).

After removing uninformative markers, 10,496 SNPs were retained for the analysis. All animals had a call rate above 95% and 99% of all SNPs had call rates higher than 99%. Missing SNPs were imputed at random based on the distribution of known

SNPs. Traits were chosen across a range of heritabilities, type (weight, behavioural or physiological) and number of QTL (Table 4.1), based on VALDAR *et al.* (2006a, suppl.; 2006b). The weight traits included in the analysis were body weight at the start of the test at six weeks of age (W6) and body weight at the end of the test at ten weeks of age (W10). Behavioural traits included three measurements. One measurement was recorded as part from an open field test (a model of anxiety) at six weeks of age, namely total activity, measured as distance travelled in a time span of five minutes (TA). Two measurements were recorded as part of a cue conditioning test at seven weeks of age, whereby freezing to a tone after association with a foot shock was measured: time spent freezing during cue in minutes (TF) and number of fecal boli after cue (FB). Physiological traits were hematocrit percentage in blood as part of a full blood count test (HC) and insulin level at 75 minutes after intraperitoneal injection with glucose dose as part of a test to model type 2 diabetes mellitus, at nine weeks of age (I75).

Table 4.1: Description of traits.

Trait	Type	Count	h^{2} [a]	QTL[a]	T [ab]
Weight at week 6 (W6)	Weight	1916	0.74	19	$x^{1/3}$
Weight at week 10 (W10)	Weight	1880	0.62	20	$x^{1/3}$
Total activity in open field test (TA)	Behavioural	1879	0.34	16	x
Time freezing during cue (TF)	Behavioural	1389	0.31	1	x
Fecal boli after cue (FB)	Behavioural	1511	0.10	2	$(x+1)^{1/2}$
Hematocrit percentage (HC)	Physiological	1578	0.11	1	x^{3}
Insulin level [c] (I75)	Physiological	1701	0.13	10	$x^{1/3}$

[a] Reported by VALDAR *et al.* (2006a, suppl.; 2006b)

[b] T = transformation

[c] Measured at 75 minutes after injection of glucose

Further information regarding the biology behind these traits can be found in the study by SOLBERG *et al.* (2006). The traits were normalised using the transformation

given in VALDAR *et al.* (2006b) and subsequently multiplied or divided by appropriate factors to avoid rounding errors in the multi-marker programme. To investigate the influence of low frequencies of missing SNPs, the trait weight at 6 weeks was analysed with missing values for SNPs treated as a separate 3rd allele with a low frequency (W6m).

4.2.2 Statistical analysis

All traits were normally distributed and analysed with models using fixed effects and covariates based on the models reported by VALDAR *et al.* (2006b). Fixed effects were sex (W6, W6m, W10, TA, FB, HC, I75), year-month (W6, W6m, W10, TA), parity (W6, I75), experimenter (TA, I75), apparatus (TF) and month (I75); covariates comprised cage density (W6, W6m, W10, I75), age in days (W6, W6m, W10) and weight (HC, I75). Cage was added as a random effect for all traits. Regarding this effect it has to be pointed out that cages consisted almost solely of animals from one family. For all practical purposes cage was nested within family (average 3.1 cages per family).

Three basic groups of models were used to compare changes in variance components, predictive ability and accuracy as a result of using genomic information. One model used only polygenic effects (1), a second model used only genomic effects (2), and a third model fitted both effects (3). For models (2) and (3), seven different sub-models were considered based on the percentage of markers that was allowed to have an effect. This included a non-mixture model using 100% and six mixture models, ranging from 70%, 40%, 10%, 7.5%, 5% to 2.5% of the SNPs having an effect. In the following, these sub models will be referred to based on their mixture percentages. All analyses were performed using a Bayesian approach as implemented in the programme iBay (JANSS 2008). The basic model using polygenic effects can be described as follows:

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{X_1 b} + \boldsymbol{X_2 c} + \boldsymbol{Z u} + \boldsymbol{e}, \quad (1)$$

where $\boldsymbol{\mu}$ fits a general mean and the vectors ***b, c, u*** and ***e*** fit the fixed, cage ($\boldsymbol{c} \sim N(0, I\sigma^2_c)$), polygenic ($\boldsymbol{u} \sim N(0, A\sigma^2_u)$) and residual effects ($\boldsymbol{e} \sim N(0, I\sigma^2_e)$, respectively. ***I*** is the identity matrix and ***A*** the additive genetic relationship matrix. $\boldsymbol{X_1}$, $\boldsymbol{X_2}$ and ***Z*** are incidence matrices relating the vectors ***b, c*** and ***u*** with ***y***. This is the mixed model which is most commonly used to predict conventional breeding values in animal breeding programmes. For the model using genomic effects, model (1) was changed to a Bayesian multi-marker association model as follows:

$$\boldsymbol{y = \mu + X_1 b + X_2 c + Qas + e,} \qquad (2)$$

where ***Qas*** fits the genomic effect, with a the vector representing effects associated with marker alleles ($\boldsymbol{a} \sim N(0,1)$), ***s*** a scaling factor modelling the variance explained by each marker, whereby ***s*** is conditionally estimated as simple normally distributed regressions and can be interpreted as a standard deviation, and ***Q*** the design matrix linking alleles with markers (JANSS 2008). Priors were assigned to the scaling factor ***s*** as follows for the non mixture models:

$$\boldsymbol{s \sim TN_{>0}\,(0,\sigma^2_g),}$$

where $\boldsymbol{\sigma^2_g}$ can be interpreted approximately as the expected average fitted variance per marker and ***TN*** denotes a truncated normal distribution. For mixture models the following scaling factors ***s*** were used:

$$\boldsymbol{s} \sim \begin{cases} \boldsymbol{N\,(0,\sigma^2_{g0})} \text{ with probability } \boldsymbol{\pi_0} \\ \boldsymbol{TN_{>0}\,(0,\sigma^2_{g1})} \text{ with probability } \boldsymbol{\pi_1 = 1 - \pi_0} \end{cases}$$

where the first distribution models the markers with on average no effect at a proportion $\boldsymbol{\pi_0}$, and the second distribution models the markers that have an effect at a proportion $\boldsymbol{\pi_1}$. The proportion of markers $\boldsymbol{\pi_1}$ varied across mixture models ranging from 100% to 2.5%. Variances for the first distribution were set to 1% of the phenotypic variance of the trait divided by the number of markers. No polygenic

effect was present and all other effects were as described for model (1). The last model, which combined both model (1) and (2), can be as described as follows:

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{X_1 b} + \boldsymbol{X_2 c} + \boldsymbol{Qas} + \boldsymbol{Zu} + \boldsymbol{e}, \quad (3)$$

where the effects are as defined earlier. Here the polygenic variance of $\boldsymbol{u}$ accounts for genetic variation which could not be explained by the genomic markers $\boldsymbol{a}$.

Estimates for the variance due to polygenic effects (σ^2_u), variance due to genomic effects (σ^2_a), cage variance (σ^2_c), residual variance (σ^2_e) and total phenotypic variance (σ^2_p) were calculated using information from all animals that had both genomic and phenotypic information. The variance due to genomic effects is calculated as the sum of the contributions to the genetic variance from each marker, plus all possible covariances due to linkage disequilibrium, taking into account the allele frequencies. The software iBay required that animals with only phenotypic data had to be excluded from the analysis.

4.2.3 Predictive ability

Predictive ability was calculated as the Pearson's correlation between a predicted observation and the corresponding realised observation. Realised observation was calculated as the phenotype corrected for fixed effects and covariates, while the predicted observation was the estimated breeding value, similar to LEGARRA *et al.* (2008). To predict these observations, a cross validation approach was used, whereby the data set was split into a validation set and a training set. The validation set contained the animals for which the observation was to be predicted, while the training set was used to estimate the parameters for the model. Size of the training set is of importance for the estimation of accurate breeding values (GODDARD and HAYES 2009) and to ensure a sufficient size of training population, a 1:5 proportion of validation to training set was used. Only animals from families with at least two members were used to create validation sets (~ 80% of all animals). These animals were split into five groups to create five validation sets. Thus each validation set

contained ~16% of all animals. This was repeated to create ten validation sets in total. Each validation set had a corresponding training set, which contained the remaining animals with phenotypic data.

Two different routines for splitting the data were used, selection within family and selection between families. For selection within family, full sib families were split between training and validation set such that each set contained at least one animal from a family. For selection between families, families were split such that no full sib family would have animals in both sets simultaneously. As a result, for selection between families no close genetic connectedness due to full sib families was available between training and validation data. In the case of selection within family, full sibs with phenotypic data linked the breeding values of the training and validation data.

4.2.4 Accuracy

The approximate change in accuracy of model (2) compared to model (1) was estimated using the formula derived by LEGARRA *et al.* (2008). The basic formula to estimate differences in accuracies between models can be described as follows:

$$\Delta r\,(g, \hat{g}) = \Delta r\,(y, \hat{y})/H\Omega,$$

where g and $\hat{g}$ are the total genetic value of the animal and its estimate, respectively, y and $\hat{y}$ are the realised observation of the animal and its predicted observation, respectively, and H^2 and Ω^2 are calculated as σ^2_g/σ^2_p and $\sigma^2_{\hat{g}}/(\sigma^2_{\hat{g}}+\sigma^2_{\hat{c}})$, respectively, with estimates for these variance components based on model (1) (LEGARRA *et al.* 2008). As with the predictive ability, this was done for selection within as well as selection between families.

4.2.5 Importance of individual markers

As indicated in Table 4.1, traits were chosen across a range of number of QTL, ranging from as low as 1 in TF and HC up to 20 in W10. This was to compare the performance of the genomic models (2) and (3) in finding regions with evidence of a marker having an increased effect, and the effect of this trait structure on the efficiency of genomic selection. Using the Bayesian approach implemented in the programme iBay (JANSS 2008), the change in odds from prior to posterior probability (PPOR) for each marker was calculated with the following formula:

$$PPOR = (\hat{p}_1 /(1- \hat{p}_1))/(\pi_1/ \pi_0),$$

where $\hat{p}_1$ is the estimate for the posterior probability of the marker having an effect, π_0 the proportion of markers with no effect and π_1 the proportion of markers that do have an effect. Results were plotted per trait for all markers, whereby a PPOR > 3.2 can be interpreted as substantial evidence for the marker to have an increased effect, a PPOR > 10 as strong evidence, and a PPOR > 100 as decisive (JANSS 2008).

4.3 Results

4.3.1 Variance components

Tables 4.2, 4.3 and 4.4 show the estimates for the total phenotypic variances, the heritabilities based on the polygenic effect, the proportions of the variance attributed to the genomic effect and the phenotypic proportions of the cage variances. Estimated variance components are based on the full data set and are presented for seven models, namely: models (1), (2) and (3), and sub-models with 10% and 2.5% of the markers assumed to be associated with an effect using models (2) and (3). Results based on sub-models using mixtures of 70%, 40%, 7.5% and 5% are not shown, because they showed the same trend as can be seen from comparing these three mixtures.

Table 4.2: Variance estimates and heritabilities for weight traits. Estimates of the total phenotypic variances (σ^2_p), variances attributed to the cage effect (σ^2_c), residual variances (σ^2_e), heritabilities based on the polygenic effect (h^2_u) and proportion of the variance attributed to the genomic effect (h^2_a) for weight traits (95%-HPD).

Trait[a]	Model	σ^2_p	σ^2_c	σ^2_e	h^2_u	h^2_a
	(1)	110.8 99.5-122.8	31.3 24.6-38.1	21.1 9.4-32.5	0.52 0.38-0.69	-
	(2) 100%	117.1 108.1-126.1	39.0 31.5-46.4	35.8 32.4-39.3	-	0.36 0.32-0.40
	(2) 10%	104.7 96.3-113.6	43.8 36.1-51.6	43.8 40.0-47.5	-	0.16 0.12-0.21
W6	(2) 2.5%	104.8 96.1-113.4	46.9 38.7-55.5	46.3 42.5-49.9	-	0.11 0.07-0.15
	(3) 100%	119.0 107.9-129.6	33.1 26.2-40.5	22.5 13.9-30.5	0.25 0.12-0.36	0.29 0.25-0.33
	(3) 10%	107.9 98.1-117.7	32.7 25.9-39.6	25.8 17.5-33.2	0.33 0.21-0.46	0.12 0.08-0.16
	(3) 2.5%	108.8 99.0-118.8	32.3 25.0-38.9	24.2 15.7-32.4	0.40 0.28-0.53	0.08 0.05-0.11
	(1)	110.0 98.7-122.0	31.5 24.3-38.3	22.2 11.2-33.8	0.51 0.36-0.66	-
	(2) 100%	118.0 108.3-126.6	39.1 31.2-46.6	36.1 32.7-39.7	-	0.36 0.33-0.40
	(2) 10%	103.7 95.1-112.3	45.9 38.3-54.2	46.1 42.3-49.8	-	0.11 0.07-0.15
W6m	(2) 2.5%	104.3 95.5-113.3	49.3 40.8-57.9	48.2 44.4-52.0	-	0.06 0.03-0.10
	(3) 100%	117.6 107.9-128.1	33.8 26.8-41.3	26.2 19.5-32.9	0.20 0.10-0.31	0.29 0.25-0.33
	(3) 10%	104.8 96.5-114.7	32.9 26.2-40.4	28.7 21.6-36.5	0.35 0.23-0.47	0.06 0.03-0.10
	(3) 2.5%	106.6 96.7-116.1	32.2 25.4-39.3	26.4 17.6-35.1	0.41 0.27-0.54	0.04 0.01-0.07
	(1)	125.7 113.6-139.9	19.0 12.8-25.4	36.6 22.1-50.3	0.55 0.40-0.71	-
	(2) 100%	133.8 124.9-143.3	28.2 21.2-35.0	48.2 43.3-52.7	-	0.43 0.40-0.47
	(2) 10%	120.2 111.5-129.3	31.6 24.2-38.6	58.9 54.0-64.5	-	0.25 0.20-0.29
W10	(2) 2.5%	119.7 111.0-128.7	34.6 27.1-42.3	62.7 57.7-68.0	-	0.19 0.14-0.23
	(3) 100%	138.3 126.8-150.9	22.3 15.9-29.0	34.1 24.1-43.9	0.22 0.10-0.33	0.37 0.33-0.41
	(3) 10%	124.9 114.3-136.1	21.2 15.0-27.4	39.0 29.6-48.6	0.32 0.20-0.45	0.19 0.15-0.25
	(3) 2.5%	125.8 113.6-137.4	20.3 14.7-26.7	37.5 25.6-47.7	0.40 0.27-0.54	0.14 0.09-0.18

[a] All traits multiplied by 10^2

Table 4.3: Variance estimates and heritabilities for behavioural traits. Estimates of the total phenotypic variances (σ^2_p), variances attributed to the cage effect (σ^2_c), residual variances (σ^2_e), heritabilities based on the polygenic effect (h^2_u) and proportion of the variance attributed to the genomic effect (h^2_a) for behavioural traits (95%-HPD).

Trait[a]	Model	σ^2_p	σ^2_c	σ^2_e	h^2_u	h^2_a
	(1)	61.1 56.3-66.1	2.3 0.2-4.5	37.6 32.0-43.0	0.35 0.23-0.46	-
	(2) 100%	60.0 56.4-63.6	3.3 1.1-5.7	40.5 37.3-43.8	-	0.27 0.24-0.30
	(2) 10%	60.7 56.6-64.8	3.8 1.3-6.1	41.7 38.4-45.1	-	0.25 0.21-0.30
TA	(2) 2.5%	60.6 56.8-64.8	5.0 2.3-7.6	43.5 40.1-46.9	-	0.20 0.16-0.24
	(3) 100%	59.7 55.6-63.7	2.5 0.1-4.5	37.8 33.5-42.3	0.11 0.03-0.19	0.22 0.18-0.25
	(3) 10%	60.7 56.4-65.3	2.5 0.2-4.5	38.4 34.2-42.5	0.12 0.04-0.20	0.21 0.15-0.26
	(3) 2.5%	61.3 56.7-66.2	2.3 0.3-4.5	37.5 32.9-42.3	0.21 0.10-0.31	0.14 0.10-0.19
	(1)	1243.4 1128.1-1351.9	45.6 2.8-94.5	790.0 655.5-924.6	0.33 0.20-0.47	-
	(2) 100%	1222.4 1132.2-1308.8	82.8 26.4-143.5	871.3 787.1-951.8	-	0.22 0.19-0.25
	(2) 10%	1206.3 1119.1-1303.4	89.2 17.1-149.6	905.7 822.5-993.2	-	0.18 0.13-0.23
TF	(2) 2.5%	1206.1 1115.6-1304.2	112.8 54.5-176.9	937.9 852.0-1024.4	-	0.13 0.08-0.17
	(3) 100%	1233.0 1133.8-1341.0	49.5 0.6-99.1	794.2 672.4-918.1	0.15 0.04-0.30	0.16 0.13-0.19
	(3) 10%	1240.0 1132.4-1353.1	47.9 1.2-97.9	809.6 683.6-929.0	0.15 0.01-0.27	0.16 0.10-0.22
	(3) 2.5%	1251.4 1147.1-1369.7	44.2 1.4-91.9	807.8 695.4-925.8	0.20 0.08-0.31	0.12 0.07-0.17
	(1)	1289.9 1191.1-1382.1	94.2 25.9-165.1	1066.2 961.3-1168.9	0.10 0.04-0.17	-
	(2) 100%	1290.0 1203.0-1382.9	105.4 40.2-170.8	1092.0 1000.8-1184.1	-	0.07 0.06-0.09
	(2) 10%	1286.0 1200.1-1383.4	111.1 49.6-177.6	1104.0 1007.0-1196.4	-	0.05 0.02-0.09
FB	(2) 2.5%	1285.0 1197.5-1383.8	116.2 53.1-182.8	1114.0 1021.1-1211.1	-	0.04 0.01-0.07
	(3) 100%	1291.0 1200.6-1385.5	92.3 28.8-160.5	1072.0 963.6-1167.8	0.05 0.00-0.10	0.05 0.04-0.07
	(3) 10%	1295.0 1202.9-1393.1	92.5 24.2-156.6	1065.0 961.4-1177.0	0.07 0.00-0.14	0.04 0.00-0.07
	(3) 2.5%	1291.0 1199.4-1393.7	97.3 29.6-165.6	1077.0 978.0-1188.2	0.06 0.01-0.13	0.03 0.00-0.06

[a] TA multiplied by 10^{-2}, TF no multiplication, FB multiplied by 10^1

Analyses based on model (1), using polygenic effects only, showed for weight traits slightly lower heritabilities compared to those reported by VALDAR *et al.* (2006b) for W10 but higher deviations for W6 and W6m (Table 4.2). The differences are likely

due to the use of slightly different fixed effects in the models for these traits here. Tables 4.3 and 4.4 showed that for behavioural and physiological traits similar heritabilities were obtained as reported by VALDAR *et al.* (2006b). Phenotypic proportions of the cage variances were low for the behavioural traits (4% to 8% of the total variance, Table 4.3) compared to the weight and physiological traits (15% to 30%, Tables 4.2 and 4.4).

Table 4.4: Variance estimates and heritabilities for physiological traits. Estimates of the total phenotypic variances (σ^2_p), variances attributed to the cage effect (σ^2_c), residual variances (σ^2_e), heritabilities based on the polygenic effect (h^2_u) and proportion of the variance attributed to the genomic effect (h^2_a) for physiological traits (95%-HPD).

Trait[a]	Model	σ^2_p	σ^2_c	σ^2_e	h^2_u	h^2_a
HC	(1)	212.0 196.7-229.8	42.3 28.4-56.5	148.0 131.3-164.1	0.10 0.01-0.19	-
	(2) 100%	211.1 196.3-227.3	44.5 32.0-58.1	152.5 139.8-165.1	-	0.07 0.05-0.08
	(2) 10%	210.6 194.0-225.8	46.3 32.4-59.8	154.5 141.2-167.6	-	0.05 0.00-0.08
	(2) 2.5%	210.5 194.7-226.5	46.8 33.2-60.9	155.8 143.7-169.6	-	0.04 0.00-0.07
	(3) 100%	213.7 196.1-229.6	40.9 27.9-54.8	145.4 128.9-162.0	0.08 0.01-0.17	0.05 0.04-0.06
	(3) 10%	212.8 196.6-229.8	41.6 27.9-55.3	147.2 130.1-162.7	0.08 0.00-0.18	0.03 0.00-0.06
	(3) 2.5%	212.9 197.0-229.7	41.0 27.7-54.7	146.3 130.1-162.3	0.10 0.00-0.18	0.02 0.00-0.05
I75	(1)	806.8 743.4-873.9	201.9 150.4-261.2	475.4 413.6-534.5	0.16 0.07-0.26	-
	(2) 100%	811.5 753.5-876.4	215.4 162.3-272.0	502.4 461.6-547.2	-	0.12 0.09-0.14
	(2) 10%	798.9 737.0-860.1	225.1 169.7-279.2	517.7 476.1-562.3	-	0.07 0.03-0.11
	(2) 2.5%	798.4 736.6-858.0	231.0 174.1-285.9	525.4 485.3-569.4	-	0.05 0.02-0.08
	(3) 100%	814.9 751.4-878.1	199.6 147.2-257.5	474.1 419.6-529.8	0.08 0.02-0.17	0.09 0.07-0.11
	(3) 10%	803.6 739.9-867.1	204.2 148.9-262.2	487.6 429.5-544.9	0.08 0.00-0.17	0.06 0.02-0.10
	(3) 2.5%	806.4 741.2-869.7	197.9 143.2-251.3	474.6 413.5-535.3	0.13 0.03-0.22	0.04 0.01-0.07

[a] HC multiplied by 10^{-2}, I75 multiplied by 10^2

Proportions of variance of weight traits, behavioural traits and physiological traits due to genomic effects (model (2)) were 22% to 31%, 22% to 33% and 25% to 30% lower, respectively, than those using the polygenic model (1) (Tables 4.2 to 4.4).

This was compensated for by an increase in cage effect and/or error effect depending on trait. Using mixtures, the proportion of the variance due to genomic effects decreased by 65% to 79%, 43% to 61% and 60% to 69% for weight traits, behavioural traits and physiological traits, respectively for the 2.5% mixture compared to model (1). The underestimation of variances due to genomic effects compared to variances due to polygenic effects may be due to incomplete LD between SNPs markers and the causal variant, worsened potentially by low frequencies of these causal variants (YANG *et al.* 2010).

In model (3), the polygenic effect essentially captured part of the genetic variance that was not accounted for by the genomic effects, and the total variance attributed to genetic effects was similar to the polygenic variances found in model (1). Proportions of the variance due to the genomic effect using model (3) were consistently slightly lower than in model (2); the proportions of the variance due to the genomic effect were 33% to 44%, 37% to 52% and 44% to 50% lower for weight traits, behavioural traits and physiological traits, respectively, than proportions of the variance due to the polygenic effects model (1). For these traits the proportions of the variance due to the polygenic effects accounted for 40% to 48%, 31% to 50% and 50% to 80%, respectively, of the heritability estimated using model (1). Mixtures showed a similar trend in model (3) as in model (2), with a 75% to 85%, 60% to 70% and 75% to 80% decrease in proportion of the variance due to the genomic effects compared to the proportions of the variance due to the polygenic effects from model (1) for the respective traits. The proportions of the variance due to the polygenic effects for weight traits, behavioural traits and physiological traits accounted for 73% to 77%, 60% to 61% and 81% to 100%, respectively, of the proportions of the variance due to the polygenic effects from model (1).

In general, a decrease in mixture, with fewer markers allowed having an effect, led to a decrease in proportions of the variance due to the genomic effects and an increase in the proportions of the variance due to the polygenic effects in all traits. The differences between mixtures however were small and, except for the weight traits,

seldom significant. Comparing W6 and W6m, treating missing alleles as a separate 3rd allele resulted in small changes in proportions of the variance due to the genomic effects with lower mixture percentages.

4.3.2 Predictive ability

Tables 4.5 and 4.6 show the average predictive ability for selection within family (W) or between families (B). Predictive abilities were calculated using ten training and validation sets, and are shown for all 3 models and their sub-model using different mixtures.

Table 4.5: Predictive abilities for selection within (W) or between (B) families for weight traits.

	W6		W6m		W10	
Model	**W** [a]	**B** [b]	**W** [a]	**B** [b]	**W** [a]	**B** [c]
(1)	0.62	0.15	0.62	0.15	0.53	0.19
(2) 100%	0.63	0.24	0.63	0.23	0.57	0.29
(2) 70%	0.65	0.26	0.65	0.26	0.58	0.31
(2) 40%	0.65	0.27	0.65	0.27	0.59	0.32
(2) 10%	0.64	0.24	0.64	0.25	0.58	0.33
(2) 7.5%	0.64	0.24	0.64	0.24	0.58	0.33
(2) 5%	0.64	0.22	0.64	0.23	0.57	0.31
(2) 2.5%	0.63	0.20	0.63	0.20	0.56	0.31
(3) 100%	0.64	0.25	0.64	0.25	0.58	0.31
(3) 70%	0.65	0.27	0.65	0.27	0.59	0.33
(3) 40%	0.65	0.27	0.65	0.27	0.59	0.34
(3) 10%	0.65	0.27	0.64	0.25	0.59	0.34
(3) 7.5%	0.65	0.26	0.64	0.25	0.59	0.34
(3) 5%	0.65	0.25	0.64	0.24	0.58	0.33
(3) 2.5%	0.64	0.24	0.63	0.23	0.57	0.31

[a] all s.e. ≤ 0.01; [b] all s.e. ≤ 0.03; [c] all s.e. ≤ 0.04

Comparing selection within and between families, it is noticeable that within family selection always performed substantially better than between families, as was expected due to the higher connectedness between the training and validation data set. When selection was carried out within family, all models performed similar in PA for most traits with only W6 and TA showing an increase of models (2) and (3) compared to (1). In contrast, using between family selection, model (1) resulted in substantially lower PA than models (2) and (3) for most traits. This was especially visible for traits with moderate to high heritabilities (e.g. W6: model (1) 0.15 vs. model (2) 100% 0.24 or TF: model (1) -0.04 vs. model (2) 100% 0.19; Tables 4.5 and 4.6). For traits with low heritabilities there was little difference in PA between model (1) and the other two models.

Table 4.6: Predictive abilities for selection within (W) or between (B) families for behavioural and physiological traits.

	TA		TF		FB		HC		I75	
Model	**W** [a]	**B** [c]	**W** [b]	**B** [d]	**W** [b]	**B** [d]	**W** [a]	**B** [b]	**W** [a]	**B** [c]
(1)	0.37	0.16	0.29	-0.04	0.21	0.10	0.33	0.08	0.42	0.08
(2) 100%	0.43	0.34	0.31	0.19	0.22	0.11	0.33	0.05	0.42	0.13
(2) 70%	0.43	0.34	0.31	0.18	0.22	0.12	0.33	0.05	0.42	0.13
(2) 40%	0.43	0.35	0.32	0.19	0.22	0.11	0.33	0.05	0.42	0.13
(2) 10%	0.42	0.34	0.33	0.20	0.21	0.11	0.33	0.06	0.42	0.14
(2) 7.5%	0.42	0.33	0.32	0.20	0.21	0.11	0.33	0.06	0.42	0.14
(2) 5%	0.41	0.30	0.32	0.19	0.21	0.11	0.33	0.06	0.42	0.13
(2) 2.5%	0.40	0.27	0.31	0.18	0.20	0.10	0.33	0.06	0.42	0.12
(3) 100%	0.43	0.33	0.33	0.17	0.22	0.13	0.33	0.06	0.43	0.13
(3) 70%	0.43	0.34	0.33	0.17	0.22	0.12	0.33	0.05	0.43	0.13
(3) 40%	0.43	0.34	0.33	0.17	0.22	0.12	0.33	0.05	0.43	0.13
(3) 10%	0.42	0.33	0.34	0.19	0.22	0.12	0.33	0.06	0.43	0.13
(3) 7.5%	0.42	0.32	0.34	0.18	0.22	0.12	0.33	0.07	0.43	0.13
(3) 5%	0.42	0.30	0.34	0.17	0.22	0.12	0.33	0.07	0.43	0.14
(3) 2.5%	0.41	0.28	0.33	0.16	0.21	0.12	0.33	0.08	0.43	0.12

[a] all s.e. ≤ 0.01; [b] all s.e. ≤ 0.02; [c] all s.e. ≤ 0.03; [d] all s.e. ≤ 0.04

Comparing different mixtures, TA was the only trait to show a significant decrease in PA for selection within as well as between families for model (2). With lower mixtures, the PA was stable at first, but with mixtures below 7.5% the PA substantially decreased compared to its highest value (0.27 vs. 0.35). W6 showed a similar pattern for selection between families with a drop-off for mixtures below 7.5% (from 0.27 to 0.20). Both W6 and W10 showed a trend for a decrease in PA for selection within family for mixtures below 7.5% in model (2). TA was the only trait to show a tendency for a lower PA for model (3) for both selection within (from 0.43 to 0.41) and between family (from 0.34 to 0.28). All other traits showed no significant decrease in PA with lower mixture percentages. As with the variance components, W6 and W6m showed no difference.

4.3.3 Accuracy

Tables 4.7 and 4.8 show the approximate increase in accuracy for model (2) compared to model (1) for all traits and mixtures. For selection within family, the increase was between 0.00 and 0.11 for all traits, except -0.01 for model (2) 100% in trait I75 and -0.03 for model (2) 2.5% in FB. For selection between families the gains were generally larger, up to 0.44 for TF, except for FB and HC, where the gains were equal or even lower than using within family selection.

Across mixtures, most traits had a tendency to show a reduced gain in accuracy when a lower percentage was used, but there appeared to be an optimum gain in accuracy for some traits. Both weight traits and TA showed an optimum with mixtures around 40%, though this was only significant when selection was within family, and TF and I75 showed a similar trend.

Table 4.7: Estimated differences in accuracies for weight traits. Estimated increase of accuracy of model (2) compared to model (1) for selection within (W) or between (B) families for weight traits.

	W6		W6m		W10	
Model	**W** [a]	**B** [c]	**W** [a]	**B** [d]	**W** [a]	**B** [b]
(2) 100%	0.01	0.15	0.01	0.15	0.06	0.16
(2) 70%	0.04	0.19	0.04	0.20	0.07	0.19
(2) 40%	0.04	0.20	0.04	0.20	0.08	0.20
(2) 10%	0.04	0.16	0.03	0.18	0.07	0.21
(2) 7.5%	0.03	0.15	0.03	0.16	0.06	0.21
(2) 5%	0.03	0.12	0.02	0.15	0.05	0.19
(2) 2.5%	0.02	0.09	0.01	0.09	0.04	0.18

[a] all s.e. ≤ 0.01; [b] all s.e. ≤ 0.05; [c] all s.e. ≤ 0.06; [d] all s.e. ≤ 0.07

Table 4.8: Estimated differences in accuracies for behavioural and physiological traits. Estimated increase of accuracy of model (2) compared to model (1) for selection within (W) or between (B) families for behavioural and physiological traits.

	TA		TF		FB		HC		I75	
Model	**W** [a]	**B** [d]	**W** [c]	**B** [d]	**W** [c]	**B** [e]	**W** [b]	**B** [e]	**W** [a]	**B** [e]
(2) 100%	0.11	0.32	0.05	0.42	0.05	0.03	0.01	-0.16	-0.01	0.19
(2) 70%	0.11	0.32	0.05	0.41	0.04	0.05	0.00	-0.17	0.00	0.21
(2) 40%	0.11	0.33	0.06	0.43	0.03	0.03	0.00	-0.17	0.00	0.20
(2) 10%	0.10	0.31	0.07	0.44	0.02	0.02	0.01	-0.09	0.00	0.23
(2) 7.5%	0.09	0.29	0.07	0.44	0.01	0.02	0.01	-0.09	0.00	0.23
(2) 5%	0.08	0.25	0.05	0.42	0.00	0.03	0.00	-0.11	0.00	0.22
(2) 2.5%	0.05	0.20	0.04	0.40	-0.03	-0.02	-0.01	-0.09	0.00	0.18

[a] all s.e. ≤ 0.01; [b] all s.e. ≤ 0.02; [c] all s.e. ≤ 0.03; [d] all s.e. ≤ 0.06; [e] all s.e. ≤ 0.10

4.3.4 Individual markers

Table 4.9 gives an overview of the number of markers showing evidence for an increased effect as identified by the change in odds from prior to posterior

probability (PPOR) in the 2.5% mixture. Note that markers do not equal QTL here, due to the fact that the effect of a QTL may be spread over several markers in a region, whereby each individual marker picks up part of the effect of the QTL. The table lists markers with a 3.2 < PPOR ≤ 10 (substantial), 10 < PPOR ≤ 100 (strong) or PPOR > 100 (decisive). Figures 4.1 and 4.2 show Manhattan plots of the PPOR per marker for model (2) and model (3), respectively.

Table 4.9: Number of markers showing varying levels of evidence of an effect. Markers are identified by the changes in odds from prior to posterior probability (PPOR) using models (2) and (3) with the 2.5% mixture. 3.2 < PPOR ≤ 10 denotes substantial evidence, 10 < PPOR ≤ 100 strong evidence and PPOR > 100 decisive evidence.

	Model 2			Model 3		
Trait	**Substantial**	**Strong**	**Decisive**	**Substantial**	**Strong**	**Decisive**
W6	26	3	1	12	6	0
W6m	5	5	0	1	5	0
W10	31	18	0	24	6	2
TA	41	13	3	39	8	0
TF	17	4	0	13	3	0
FB	5	2	0	3	1	0
HC	9	1	0	4	0	0
I75	6	2	0	1	2	0

Generally model (2) detected more markers per category and trait than model (3), with one exception, namely for the number of markers with decisive evidence in the weight trait W10. Based on model (2), the two weight traits and TA showed the highest numbers (30 and 57 in total), followed by TF (21 in total). The three traits with the lowest heritabilities, FB, HC and I75, showed the lowest numbers of markers (ranging from 7 to 10 in total). In contrast to the variance estimates, predictive ability and accuracy, for which treating missing alleles as a separate 3rd allele did not change their estimates, the number of markers with increased evidence for an effect was much lower for W6m than for W6.

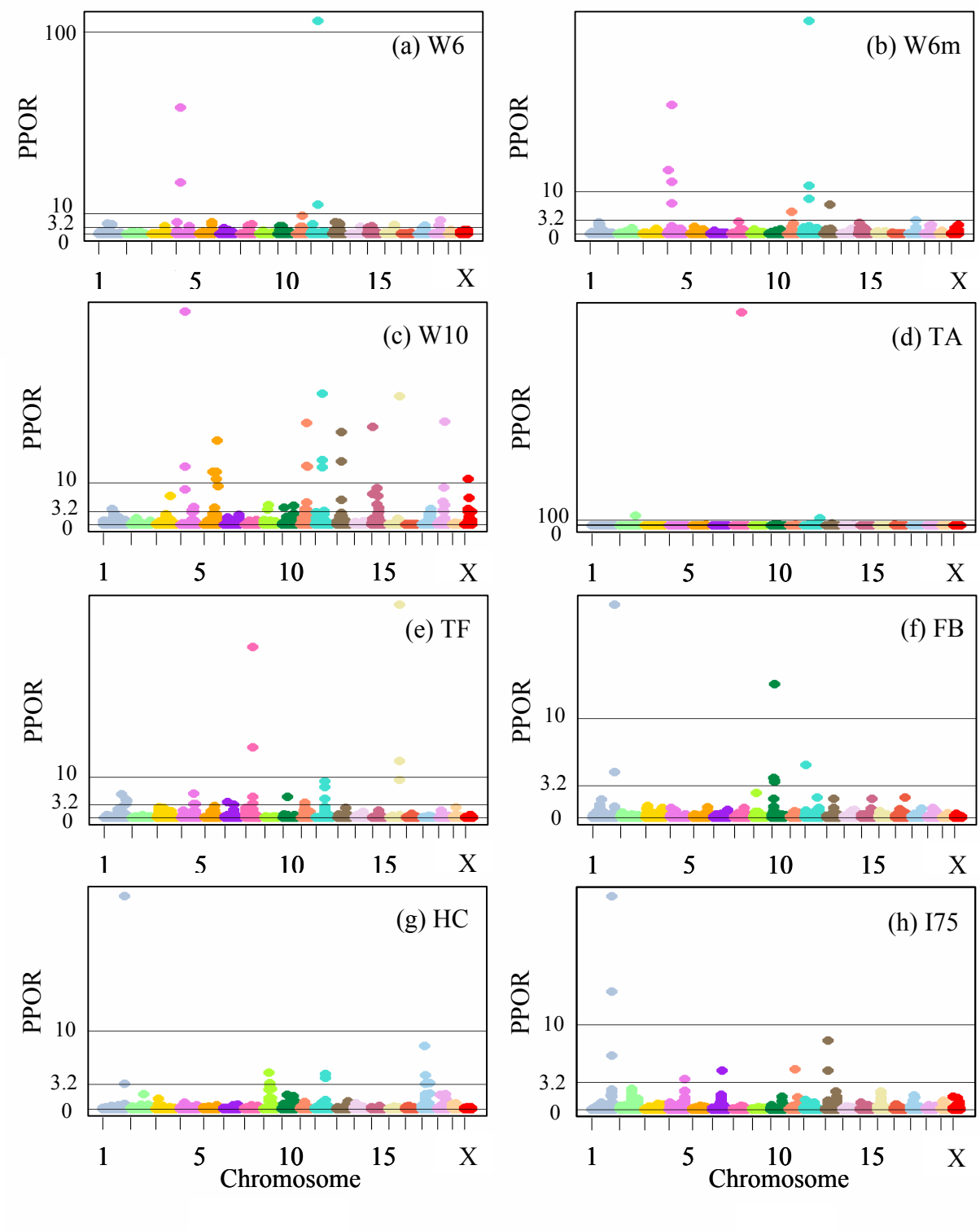


Figure 4.1: Distribution of odds ratios of SNP effects based on model (2). Markers are identified by the changes in odds from prior to posterior probability (PPOR) using model (2) with the 2.5% mixture. 3.2 < PPOR ≤ 10 denotes substantial evidence, 10 < PPOR ≤ 100 strong evidence and PPOR > 100 decisive evidence.

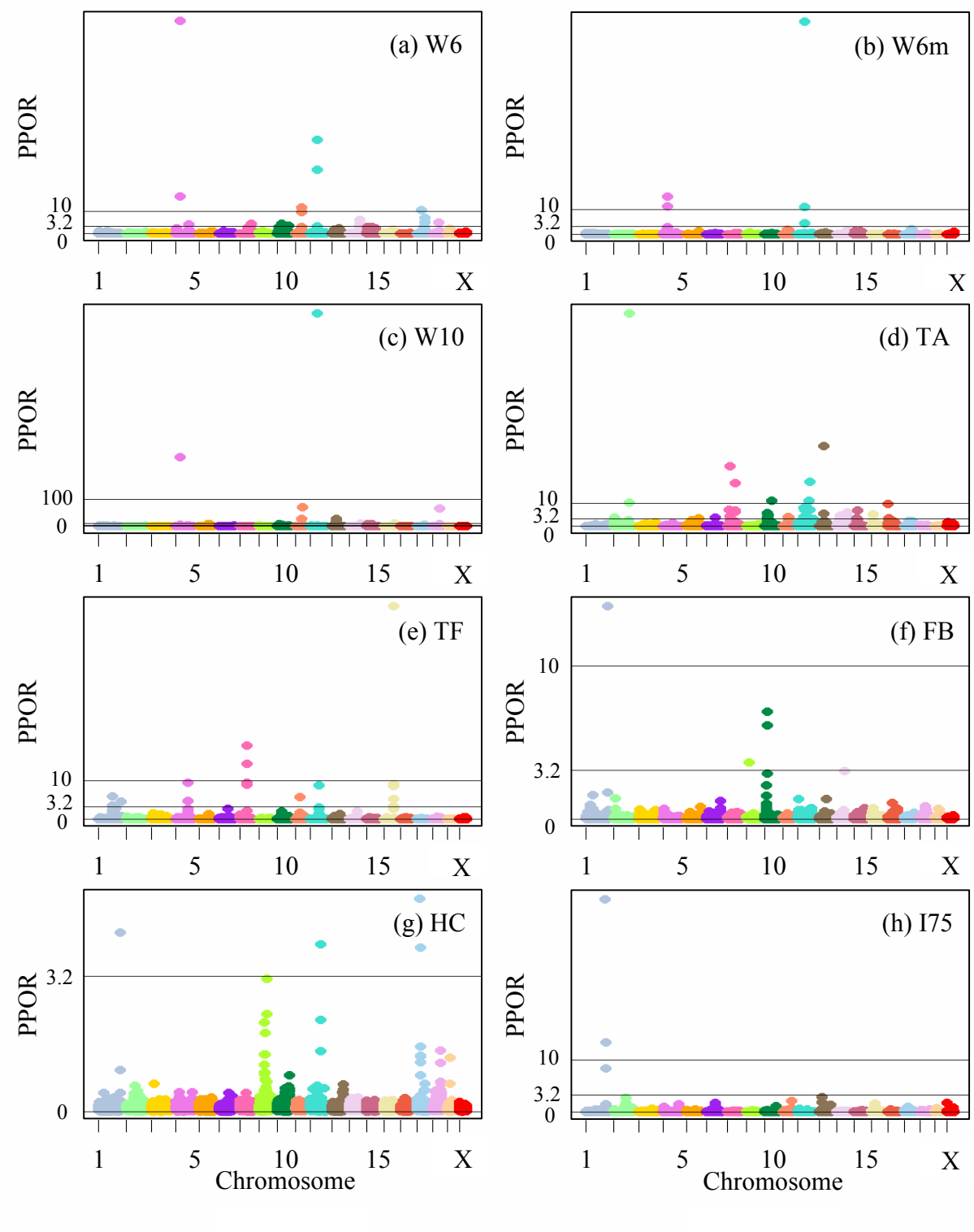


Figure 4.2: Distribution of odds ratios of SNP effects based on model (3). Markers are identified by the changes in odds from prior to posterior probability (PPOR) using model (3) with the 2.5% mixture. 3.2 < PPOR ≤ 10 denotes substantial evidence, 10 < PPOR ≤ 100 strong evidence and PPOR > 100 decisive evidence.

The figures show that the patterns between the two models were similar in terms of the chromosome on which the regions are located. Some variation was visible in the relative weight of markers located closely together due to the fact that markers situated near a QTL might each pick up part of the QTL effect. An example can be seen on chromosome 11 for trait W6, where for model (2) two adjacent markers showed a PPOR of 105 and 14, respectively, while the same markers for model (3) showed a PPOR of 29 and 42, respectively.

4.4 Discussion

4.4.1 Heritabilities

In general, higher heritabilities resulted in an increase in predictive abilities of genomic selection for all traits. Similar results were found for a different set of traits from this data set (LEGARRA *et al.* 2008; USAI *et al.* 2009), with predictive abilities as high as 0.67 for a trait with a high heritability (weight, $h^2 = 0.74$), but as low 0.27 for a trait with a low heritability (body length, $h^2 = 0.13$). However, the relationship between heritability and predictive ability was far from linear, as can be seen when comparing for example TF and I75, where the latter had a lower heritability but a higher PA when selection was within family. A similar lack of linear association with heritability was seen for the increase of accuracy of the genomic model over the polygenic model. Traits with moderate heritabilities showed larger increases in accuracy than traits with high heritabilities, but the three traits with low heritabilities did not follow this trend. This might indicate that other factors besides the heritability have an influence on the predictive ability of a model.

4.4.2 QTL and individual marker distribution

In addition to the heritability, the influence of QTL distribution on the predictive ability and accuracy of a trait was also investigated. As pointed out earlier, individual markers found to have an increased chance to have an effect do not equal QTL, and for most traits more markers were found to have an substantial effect than QTL found by VALDAR *et al.* (2006a). Across the traits in this study, the number of

markers depended partially on the heritability, but especially for the traits with low to moderate heritabilities the number of markers varied substantially independent of the magnitude of the heritabilities.

A clear tendency for traits with fewer markers to have a lower PA was seen when selection was between families. The only exception was HC, which had the lowest PA but not the lowest total number of markers. However, this trait had a relatively high number of markers classified as 'only' substantial compared to the other traits and few strong or decisive markers. Selection within family showed the aforementioned trend to a lesser extent.

Simulation studies, e.g. by ZHONG *et al.* (2009) and KIZILKAYA *et al.* (2009), have shown that the number of QTL affecting a trait influences the performance of genomic selection, though the influence differed depending on the method that was used. The latter study, by KIZILKAYA *et al.* (2009), found that an increase of the number of QTL explaining a set variance of a trait, which meant that less variance was attributed to a single QTL, led to a decrease in correlations between true and predicted genotype in both purebred (from 0.39 to 0.20) and multi breed (from 0.42 to 0.30) situations using genomic information only.

The number of markers was also of influence on the increase in accuracy, mainly visible in traits with low numbers of QTL (as given in Table 4.1). The traits HC and FB, with few QTL, showed little to no gain in accuracy and in some cases even a loss. TF is an interesting trait, because it had both a moderate number of QTL as well as a moderate heritability, but the largest gain in accuracy of all traits when selection was between families.

4.4.3 Behavioural traits versus weight traits and physiological traits

Analysis of the variance components indicated that behavioural traits showed in general a much lower cage effect (4% to 7% for the polygenic model) than other traits (15% to 28%), which was the case for models 2 and 3 as well. This higher

environmental effect for the weight traits and physiological traits was also found by VALDAR *et al.* (2006b) and various reasons, such as the more automated process used to record behavioural phenotypes, were thought to be the reason. Apart from the difference in cage effect however, type of trait did not show a clear influence on other parameters such as predictive ability and accuracy. Behavioural traits are generally difficult to collect in large quantities and difficult to measure directly and therefore require a suitable proxy. Research has found SNPs related to aggressive behaviour in dogs (VAGE and LINGAAS 2008), but little research has been done into the efficiency of genomic selection in behavioural traits.

4.4.4 Selection within or between families

When selection was within family, all three models performed similar. If extensive pedigree information was available, genomic information provided only a small benefit over polygenic selection only. However, as soon as family ties were less close, as with between families selection, genomic information became a lot more valuable, as was found in other studies (LEE *et al.* 2008; LEGARRA *et al.* 2008). This change was to some extent dependent on some of the factors discussed before, namely the heritability and number of QTL. For FB and HC, two traits with low heritabilities and few QTL, genomic selection did not lead to an increase in PA, and led to low or even negative increases in accuracy. Compare this to TF, a trait with a moderate heritability despite few QTL, where genomic information led to an increase in PA and a substantial increase in accuracy when selection was between families. In I75, a trait with more QTL than the three aforementioned traits but a low heritability, inclusion of genomic information led to an increase in PA and accuracy for selection between families.

4.4.5 Inclusion of a polygenic effect

Adding a polygenic effect to the genomic effects model mainly influenced the variance estimates (by picking up the part of the genetic variance that was not captured by the genomic variance) and the number of QTL to be found, but had little influence on the PA. Both LEGARRA *et al.* (2008) and DE LOS CAMPOS *et al.* (2009)

found an increased PA in this data set when using genomic information instead of polygenic information, but little difference between a genomic model or a combined model. A simulation study by CALUS and VEERKAMP (2007) showed slight increases of accuracy when adding a polygenic effect, but this was dependent on the linkage disequilibrium between adjacent markers. The same study also showed that genomic selection underestimated the genetic variance, but that this was improved by adding a polygenic component, as was seen in this study.

4.4.6 Structure of the data set

Regarding the structure of the data set, imputation of missing SNPs was evaluated in the trait W6. Treating missing SNPs as a 3rd allele affected the discovery of markers with increased evidence of having an effect, but this difference was mainly due to a large reduction in markers classified 'only as substantial'. This reduction had no influence on other aspects of the analysis, with PA and accuracy being practically identical between W6 and W6m.

4.4.7 Influence of proportion of markers

Reducing the numbers of markers that was allowed to have an effect influenced the variance estimates, but had no significant effect on the PA and accuracy for most traits. Mixture models catch less of the variance attributed to genomic effects, but give better estimates of the single SNP effects. As a consequence, this may lead to similar accuracies of prediction. TA was the only trait to show a significant decrease in PA for selection within as well as between family, and not until mixtures reached a percentage below 7.5%. The weight traits showed a similar trend for some models, but for most traits the mixtures indicated that even with large decreases down to 2.5%, no change in PA occurred. Even though estimates for the PA were by and large not significantly different from each other, they showed a trend for an optimum increase at certain mixture percentages, with highest values often around mixtures of 40%.

SU *et al.* (2009) found similar results in dairy cattle when looking at the squared correlation between breeding values in bulls across a range of percentages and traits. Reducing the percentages eventually led to lower correlations, but, depending on trait, the decline was small and did not appear until percentages were below 20% (e.g. in the trait fat percentage in milk). In traits with a limited number of large QTL a larger part of the variance is accounted for by these QTL than in traits where no large QTL are present and variance is shared more uniformly between smaller QTL. Reducing the number of SNPs might lead to an even higher peak in terms of variance explained by few, large QTL, as was shown by SU *et al.* (2009, Figure 2). This may be a reason for a slightly lower PA when the variance is distributed less evenly, which could be seen when comparing traits with more QTL (e.g. TA) to traits with few QTL (e.g. FB).

Due to the large costs of genotyping, reductions of these costs through either low-density SNP panels or methodologies that reduce the numbers of animals to be genotyped are of great importance. Research in genome-wide association studies has found that a two-stage design with pre-selection of SNPs between steps can reduce costs greatly without reducing the power of the study (SATAGOPAN and ELSTON 2003; LI 2008). Another strategy is the imputation of haplotypes or missing genotypes, for example long-range phasing (DAETWYLER *et al.* 2010; HICKEY *et al.* 2010). These results indicate that there might be an optimum to the proportion of SNPs to be used for genomic selection, where a high efficiency is combined with lower financial costs. Depending on trait characteristics such as heritability and QTL structure, a (pre-selected) subset of markers may be sufficient to arrive at this optimum value, and a low-density SNP panel could then be used. However, breeding programmes usually consider more than ten traits and, depending on overlap of the selected SNP markers, the total number of selected SNPs may be considerably larger than the number of SNPs selected for a single trait.

4.5 Conclusions

Genomic selection generally performed better than traditional polygenic selection, as seen by the increase in predictive ability and accuracy. It was particularly beneficial in situations where selection was across families and polygenic selection was thus not able to perform well. Larger increases in predictive ability and accuracy were found for traits with lower heritabilities, but the underlying QTL distribution had an important effect. Traits with fewer QTL also showed lower predictive abilities and in certain cases even a loss of accuracy. Behavioural traits showed a lower environmental variance than other traits, but no difference in efficiency of genomic selection compared to other traits. Models including a polygenic effect with the genomic effect captured more of the genetic variance, but did not improve the predictive ability of the models. The data set was restricted to genotyped animals only; models that can incorporate non-genotyped animals directly might show different results due to for example lower errors of estimation of fixed effects and higher accuracy of the polygenic effects.

Reducing the number of markers did not significantly change the predictive ability for most traits, particularly when selection was within family. The mixture approach showed that models using a lower percentage of SNPs performed efficiently across a large range of percentages, which may be of greater importance in the future due to the increasing sizes of SNP panels. Only at very low percentages of 7.5% and less did the predictive ability for some traits decrease. Most traits showed that a substantial reduction in the number of SNPs does not reduce the efficiency of genomic selection significantly. Depending on the trait, fewer markers would be sufficient to ensure consistently high efficiency of genomic selection, thus reducing the costs of genotyping and enabling implementation of genotyping at a larger scale.

Chapter 5 – Characterisation of the Illumina PorcineSNP60 Panel

5.1 Introduction

The aim of this chapter was to describe the results of genotyping 576 Yorkshire pigs using the PorcineSNP60 panel, to assess marker characteristics including chromosome coverage, allelic systems, minor allele frequencies, Hardy-Weinberg Equilibrium and linkage disequilibrium. The animals were genotyped as part of a study into the genetic background of aggressiveness in pigs. Behavioural traits are often difficult to measure, because they tend to be time consuming and costly. Establishing genetic markers as indicators for behavioural traits can further our understanding of these traits. Subsequent incorporation of selection for behavioural traits via genomic selection in a breeding goal has the potential to benefit animal welfare greatly. The data, described by D'EATH *et al.* (2009) and TURNER *et al.* (2009), are of major interest for research into the genomic background of behavioural traits in pigs and a more in-depth study of the characteristics of these genotype data is of great interest.

5.2 Materials and methods

5.2.1 Animals and SNPs

Purebred Yorkshire pigs from a dam line nucleus herd were used in this study. These animals were part of a behavioural study on the genetics of temperament in pigs which was conducted between October 2005 and January 2007 in Ransta, Sweden (D'EATH *et al.* 2009; TURNER *et al.* 2009). The group comprised 135 litters from 114 dams and 42 sires. DNA was collected during the study and samples were genotyped in 2009 using the PorcineSNP60 panel from Illumina. Genotyping was performed by an external contractor in a three-day process using an Illumina extraction protocol and Illumina Infinium Multi-Use Assay Protocol. Afterwards, data were available for 576 animals and 62,163 SNPs. Of these animals, 552 had a call rate (the percentage of markers assigned a genotype per animal) higher than 99%.

5.2.2 Description of characteristics

Chromosome coverage was determined as the number of SNPs as well as the average distance in base pairs between adjacent SNPs. To assess the distribution of allele frequencies, the number of SNPs per minor allele frequency (MAF) was calculated whereby frequencies were grouped in categories of 5%. Low minor allele frequencies may present difficulties due to the fact that a large sample size is required to estimate their effect reliably. As a result, SNPs with a low MAF are sometimes excluded from further analysis. Per SNP the frequencies of the alleles was counted. Using an exact test, the Hardy-Weinberg Equilibrium (HWE) state of each locus was established, as well as the expected and observed heterozygosity levels. Various causes – such as selective breeding in livestock – can lead to deviations from the equilibrium. Linkage disequilibrium (LD) was estimated as the squared correlation between two markers as proposed by HILL and ROBERTSON (1968). LD can develop, for example, as a result of mutations, drift or selective breeding. For this analysis, markers with a MAF below 1% were not considered. The measure of LD was calculated pair-wise, between adjacent SNPs as well as among SNPs situated within one million base pair (1 Mb) windows. The decay of LD over distance was evaluated whereby distances were grouped in categories of 25,000 base pairs. All analyses were performed using SAS (SAS 2002) and PLINK (PURCELL *et al.* 2007).

5.3 Results and discussion

5.3.1 Chromosome coverage

The data set consisted of 576 animals, 241 males and 335 females. Of the 62,163 SNPs in the data set, 13,970 SNPs (22%) were not assigned to a chromosome. Figure 5.1 shows the distribution of the remaining 48,193 markers over the chromosomes. SNPs were distributed throughout the genome, ranging from as few as 19 on the Y-chromosome up to 6622 on chromosome 1. The average number of SNPs per chromosome (excluding the Y-chromosome) was 2535. Figure 5.2 shows the average distance between two SNPs per chromosome, based on 48,174 markers. The average distance between adjacent SNPs over all chromosomes was 56,495 base

pairs. These distances are generally slightly larger than the distances found by RAMOS *et al.* (2009), who reported distances between 33,700 and 42,800 base pairs, except for the X chromosome (67,500 base pairs). The reasons could be that they used slightly more SNPs in their study, but also that with updated maps, positions of SNPs along the chromosomes may become known with greater precision.

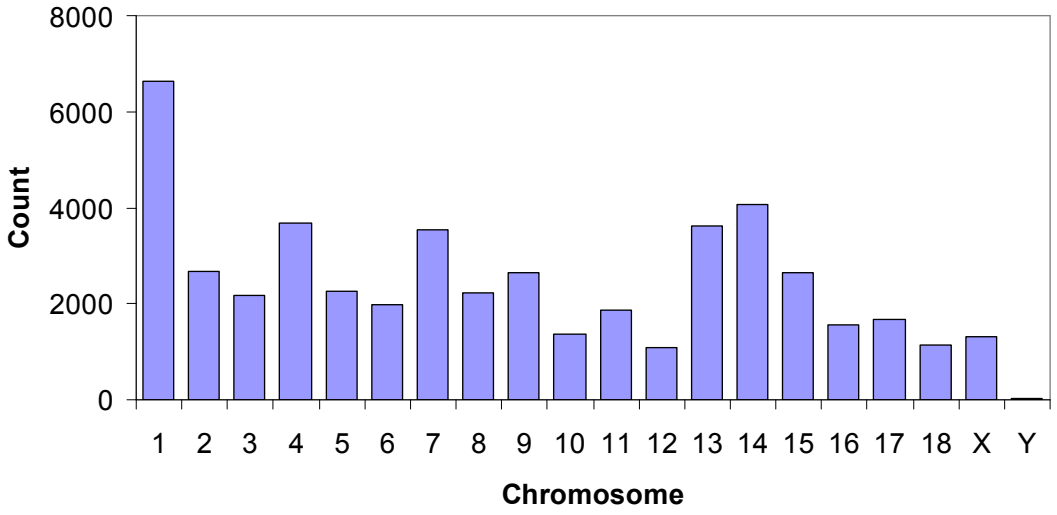


Figure 5.1: SNP distribution. Distribution of SNPs over the 18 autosomes and the two sex chromosomes.

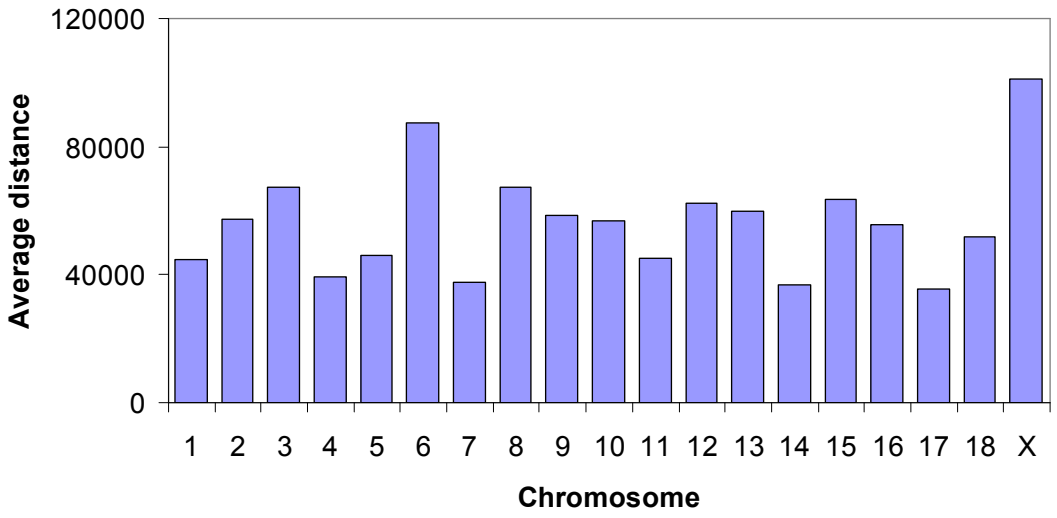


Figure 5.2: Average SNP distance. Average distance in base pairs between adjacent SNPs per chromosome.

The average distance between adjacent SNPs varied considerably among chromosomes. Chromosome 17 had the lowest average distance between SNPs at 35,594 base pairs, followed closely by chromosome 14 at 36,654 base pairs and chromosome 7 at 37,506 base pairs. The X-chromosome had the highest average distance between SNPs at 100,988 base pairs, followed by chromosome 6 at 87,307 base pairs and chromosome 8 at 67,466 base pairs. Due to its low number of SNPs, the Y-chromosome was not considered for this analysis.

5.3.2 Allelic systems

The majority of heterozygous allelic systems were formed by AG-pairs, followed closely by CT-pairs. Together these two accounted for 80% of all allelic systems. AC and GT together accounted for 18%, while AT and CG together formed the remaining 2% of all allelic systems. It is unknown what causes the difference in frequencies between pairs. However, SNPs for the panel were selected such that each allele accounts for roughly 25% of all alleles, and the fact that certain pairs show a much higher frequency than others may simply be a consequence of this choice. Figure 5.3 shows the distribution of percentages over the six different allelic systems based on heterozygous SNPs. Of the homozygous systems, C and G formed 30% each, while A and T formed 20% each.

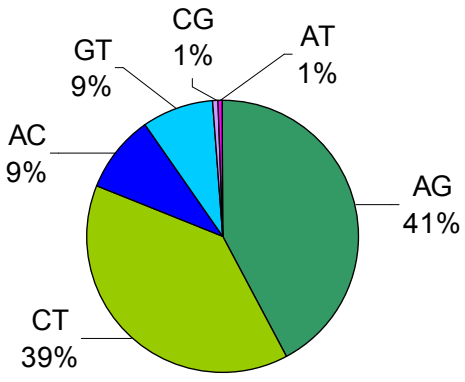


Figure 5.3: Allelic systems. Frequencies of the allelic systems AG, CT, AC, GT, CG and AT.

5.3.3 Minor allele frequency

The numbers of SNPs across MAF-categories was fairly stable, with a slight tendency for more SNPs with a higher MAF (Figure 5.4). The average MAF based on 50,976 heterozygous SNPs was 0.26 ± 0.14. This is in line with the average MAF of 0.27 for the individuals used for the design of this SNP panel (RAMOS *et al.* 2009), as well as the average MAF of 0.27 on chromosome one for a Landrace dam line using this SNP panel (HUISMAN *et al.* 2010). Per chromosome the average MAF ranged from 0.23 (chromosome 16) to 0.29 (chromosome 2). A closer look at the category of SNPs with a MAF between 0% and 5% showed that frequencies of SNPs with very low MAF were low (Table 5.1).

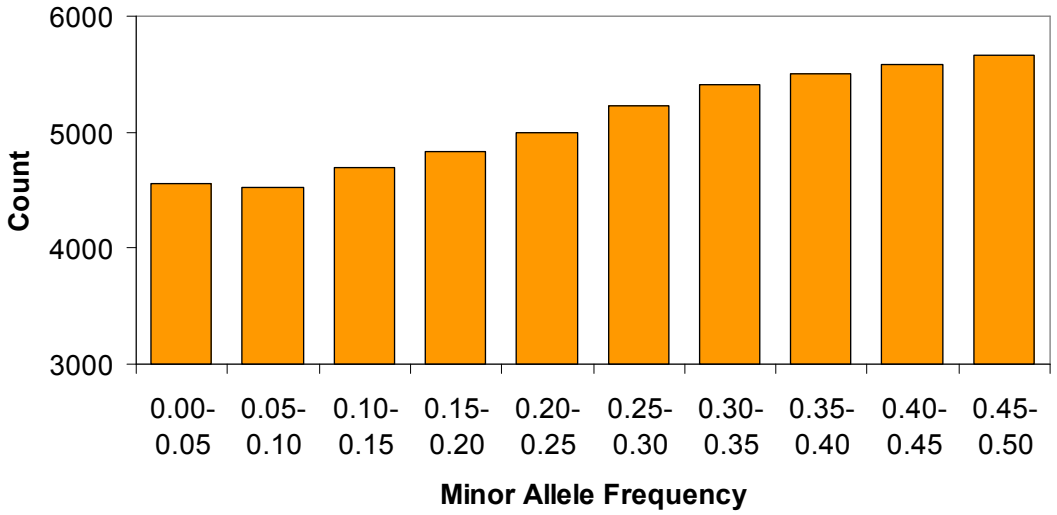


Figure 5.4: Minor allele frequencies. Number of SNPs per category of minor allele frequency (in increments of 0.05).

Table 5.1: Minor allele frequencies below 5%. Frequencies of minor alleles for MAF values of 1%, 2% and 5%.

MAF	Count	Percentage
< 0.01	1,167	2
< 0.02	2,090	4
< 0.05	4,545	9

5.3.4 Hardy-Weinberg equilibrium

The majority of SNPs were found to be in HWE (threshold at P < 0.001; Table 5.2). The observed heterozygosity in all SNPs was only slightly higher than the expected heterozygosity (0.349 versus 0.345). An average heterozygote excess of 0.004 (ranging from -0.088 to 0.077) was found for SNPs in HWE. SNPs that were not in HWE had a slightly higher expected and observed heterozygosity, as well as a heterozygote excess of 0.007 (ranging from -0.044 to 0.500), but this deviation was not significantly different from the SNPs that were in HWE (Table 5.3).

Table 5.2: Hardy-Weinberg Equilibrium. Hardy-Weinberg Equilibrium (HWE) state for informative loci (threshold at P < 0.001).

HWE	Count	Percentage
All SNPs	50,976	100
SNPs in HWE	49,166	96
SNPs not in HWE	1,810	4

Table 5.3: Heterozygosity. Observed and expected heterozygosity (s.e. as subscript).

Heterozygosity	All SNPs	SNPs in HWE	SNPs not in HWE
Expected	$0.345_{0.147}$	$0.343_{0.148}$	$0.378_{0.112}$
Observed	$0.349_{0.152}$	$0.348_{0.151}$	$0.385_{0.174}$
Deviation	$0.004_{0.031}$	$0.004_{0.021}$	$0.007_{0.125}$

5.3.5 Linkage disequilibrium

The average LD between adjacent SNPs was 0.36 ± 0.37 based on 38,059 SNP pairs, with an average distance of 52,659 base pairs. Per chromosome the LD varied considerably from 0.29 to 0.44 (Figure 5.5). Considering all SNPs within 1 Mb windows, the average LD was 0.28 ± 0.32 based on 331,666 SNP pairs in total, with an average distance of 215,618 base pairs. Figure 5.6 shows the average LD per chromosome within 1Mb windows. As with the average over all chromosomes, the LD per chromosome decreased compared to the LD presented in Figure 5.5, but the pattern across chromosomes remained constant.

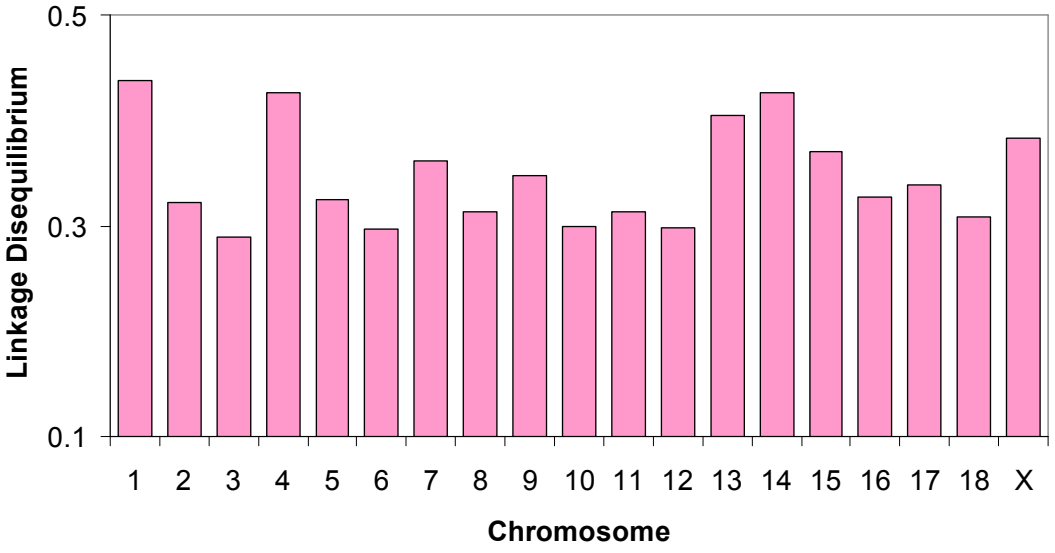


Figure 5.5: Linkage disequilibrium between adjacent SNPs. Average linkage disequilibrium (r^2) between adjacent SNPs per chromosome.

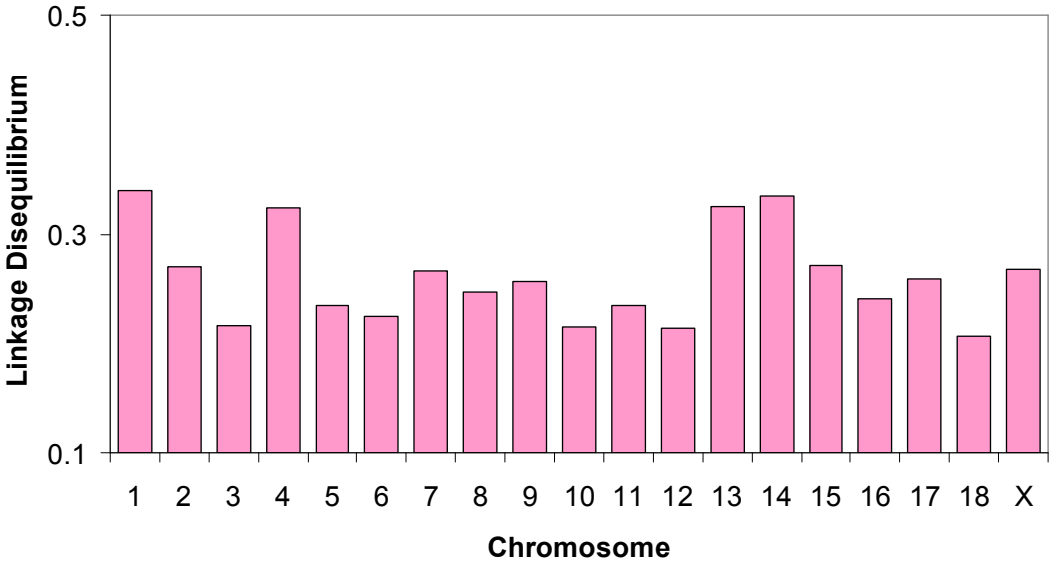


Figure 5.6: Linkage disequilibrium within 1Mb. Average linkage disequilibrium (r^2) between SNP located within 1Mb windows per chromosome.

With increasing distance between markers, the LD decreased, as was expected (Figure 5.7). HARMEGNIES *et al.* (2006) evaluated LD in two commercial pig populations (Large White and Large White x Landrace) using 34 microsatellite

markers on two chromosomes. They found significant LD extending up 40 cM in distance, with an r^2 of 0.15 to 0.50 for markers less than 1 cM apart. DU *et al.* (2007) calculated LD in commercial sire lines (Pietrain and Duroc) and four commercial dam lines (Large White and Landrace) using ~4500 autosomal SNPs. For SNPs less than 1 cM apart they found an r^2 of 0.21 to 0.51, which reduced to 0.01 at a distance of 40 cM. For chromosome one HUISMAN *et al.* (2010) found an r^2 of up to 0.27 for a Pietrain sire line and up to 0.22 for a Landrace dam line, slightly lower than generally found in pigs. In addition to differences between sire and dam lines, considerable difference in LD also exists between European and Chinese pig populations, with a much higher extent of LD in European breeds (AMARAL *et al.* 2008). This may have been due to more intensive selection, which increased the LD in European breeds, or due to higher LD in their ancestral population.

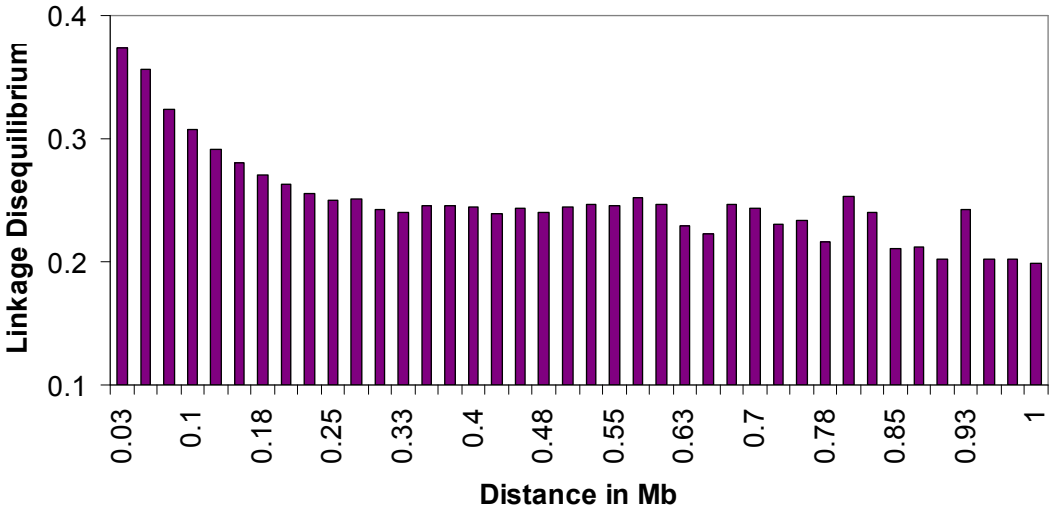


Figure 5.7: Linkage disequilibrium versus distance. Average linkage disequilibrium (r^2) versus the distance in Mb in increments of 25,000 base pairs.

The LD has been evaluated also in various other species. Compared to the mouse data set used in Chapter 4 (VALDAR *et al.* 2006a), LD in this data set was lower. Other species, e.g. cattle (BANOS and COFFEY 2010), sheep (MCRAE *et al.* 2002) and dogs (SUTTER *et al.* 2004), showed a similar pattern of high LD at short range, with a decay of LD with increase in distance, but at a lower rate than the decrease in LD often estimated in human populations (REICH *et al.* 2001).

An LD of 0.30 is generally considered to be the minimum useful for genomic selection (e.g. SARGOLZAEI *et al.* 2008; BANOS and COFFEY 2010), and was found for markers that were located within 75,000 to 100,000 base pairs of each other or less. When considering only SNPs with an LD of 0.30 or higher, the average LD was 0.67 ± 0.25, based on 109,171 SNP pairs. For 22,434 SNP pairs the LD was > 0.99, with the relative distribution of these SNP pairs over chromosomes in line with the distribution of the total number of SNPs per chromosomes (Figure 5.1). HARMEGNIES et al. (2006) also found that useful LD extended over a much shorter region than the general extent of significant LD. BANOS and COFFEY (2010) reported a shorter range of up to 50,000 base pairs for meaningful LD in cattle.

Knowledge of the genomic structure of this pig population is of major importance for genome-wide association studies into the background of behavioural traits. In addition, it provides valuable information for genomic selection. Moreover, the PorcineSNP60 panel has been used for a range of purposes recently, including paternal identification (DUIJVESTEIJN *et al.* 2010b), determination of inbreeding (SILIÓ *et al.* 2010), discovery of signatures of selection in sire and dam lines (HUISMAN *et al.* 2010) and evaluation of colonisation history (SOUZA *et al.* 2010), all of which require in-depth knowledge of the genomic structure of the pig population.

5.4 Conclusions

The SNP data set described here is valuable source of information on the genomic structure of a pig population. The extent of LD in this population was substantial, but quickly decayed with increase in distance. However, compared to other species this decay was less pronounced and meaningful LD was found along a considerable range. The comprehensive coverage of the genome as well as the extent of linkage disequilibrium in the pig genome suggest that genome-wide association studies and genomic selection may be successful to improve traits that are difficult to measure – such as behaviour characteristics – to enhance animal welfare.

Chapter 6 – The use of molecular genetic information for selection against aggressive behaviour in pigs

6.1 Introduction

Behavioural data are often costly and difficult to measure and tend to have low to moderate heritabilities. Single nucleotide polymorphism (SNP) markers associated with quantitative trait loci (QTL) of these traits may be of great benefit for their genetic improvement using marker-assisted selection (FERNANDO and GROSSMAN 1989) or genomic selection (MEUWISSEN *et al.* 2001). The aim of this chapter was twofold: firstly, the genomic structure of aggressiveness traits was assessed to investigate if markers with sufficient evidence for QTL were present. Secondly, the efficiency of genomic selection for aggressiveness traits was assessed using models with various percentages of SNP markers considered to be associated with the traits of interest, with or without inclusion of a polygenic effect.

6.2 Material and methods

6.2.1 Animals and SNPs

Data were available on 1657 pigs, purebred Yorkshire and crossbred Yorkshire x Landrace, from a dam line nucleus herd. The animals were part of a behavioural study into the genetics of temperament in pigs which was conducted between October 2005 and January 2007 in Ransta, Sweden (D'EATH *et al.* 2009; TURNER *et al.* 2009). The animals comprised of 322 litters from 250 dams and 85 sires, with a pedigree that consisted of 2419 animals in total. Lesion scores, defined as the number of fresh lesions counted by a single observer and judged subjectively based on colour and age of scabbing, were determined, whereby no weight was given to the size of the lesion. Six lesion scores traits were recorded: lesion scores at the anterior (LSA1), central (LSC1) and caudal (LSE1) region of the body at mixing and lesion scores at the same locations (LSA2, LSC2 and LSE2, respectively) at three weeks post mixing. Behaviour of the animals was recorded in the first 24 hours after mixing and comprised of duration of reciprocal aggression (RA), duration of delivery of non-reciprocal aggression (DNRA) and duration of receipt of non-reciprocal aggression (RNRA) in seconds per pig (Table 6.1). Reciprocal aggression was

defined as fights which lasted ≥ 1 s and saw both pigs to be pushing, head knocking or biting the opponent. Delivery or receipt of NRA were recorded when one pig received aggression without retaliating (TURNER *et al.* 2009).

Table 6.1: Description of traits. Description of lesion scores at mixing, lesion scores post mixing and behavioural traits, number of observations (N), means after transformation (s.d. as subscript) and heritabilities (s.e. as subscript).

Trait	N	Mean [a]	h^{2} [b]
Lesion Score at Mixing:			
Anterior region (LSA1)	1657	257 $_{109.2}$	0.26 $_{0.02}$
Central region (LSC1)	1657	209 $_{111.3}$	0.25 $_{0.03}$
Caudal region (LSE1)	1657	141 $_{103.2}$	0.21 $_{0.02}$
Lesion Score Post Mixing:			
Anterior region (LSA2)	1655	230 $_{54.8}$	0.43 $_{0.04}$
Central region (LSC2)	1655	227 $_{58.9}$	0.35 $_{0.03}$
Caudal region (LSE2)	1655	149 $_{69.7}$	0.19 $_{0.02}$
Behavioural traits:			
Reciprocal Aggression (RA)	1181	539 $_{199.7}$	0.43 $_{0.04}$
Delivery of Non-RA (DNRA)	1181	316 $_{143.6}$	0.31 $_{0.04}$
Receipt of Non-RA (RNRA)	1181	277 $_{196.8}$	0.08 $_{0.03}$

[a] transformation $Y = \log_e (1 + \text{observation})$ according to TURNER *et al.* (2009)

[b] heritability as estimated by TURNER *et al.* (2009)

DNA was available from a subset of 552 purebred Yorkshire pigs, which were genotyped in 2009 using the PorcineSNP60 panel from Illumina. The animals comprised 135 litters from 114 dams and 42 sires with on average 4.1 piglets per fullsib family. For an in-depth description of the marker data set, see Chapter 5. After removing uninformative markers, 50,203 markers were used for further analysis. Genotypes were retained regardless of their minor allele frequency. For all animals,

at least 99% of all SNPs was known. Based on the similar results for efficiency of genomic selection with missing SNPs imputed at random or treated as a 3rd allele (W6 vs. W6m in Chapter 4), the decision was made not to impute markers in this data set. The average linkage disequilibrium between adjacent SNPs, estimated as the squared correlation between two markers as proposed by HILL and ROBERTSON (1968), was 0.36 ± 0.37 (Chapter 5.3.5).

6.2.2 Statistical Analysis

All traits were normally distributed and analysed with models using fixed effects and covariates based on the models reported by TURNER *et al*. (2009). Fixed effects included in the model were sex (male or female), line (purebred or crossbred – only for the entire data set using all phenotyped animals) and batch at mixing (14 different mixing days). Weight at mixing was included in the model as a covariate and pen in which the animals were mixed as a random effect.

Three basic groups of models were used to compare changes in variance components, predictive ability and accuracy as a result of using genomic information. The first model used only polygenic effects (1), the second model used only genomic effects (2), and the third model fitted both effects (3). For models (2) and (3), five different sub-models were evaluated based on the percentage of markers considered to have an effect on the trait. This included a non-mixture model assuming all SNPs have an effect on the trait (100%) and four mixture models, ranging from 25%, 10%, 2.5% and 1% of SNPs considered having an effect on the trait. In this chapter, these sub-models will be referred to based on their mixture percentages. All analyses were performed using a Bayesian approach as implemented in the programme iBay (JANSS 2008). For a detailed description of these models, see Chapter 4.2.1.

Based on model (1) and using the phenotypes of all animals, their pre-adjusted phenotypes were estimated as the observed phenotype minus the corresponding estimates of the fixed effects and covariates. Overall, each trait was analysed three

times, for all phenotyped animals (using model (1) only), for the genotyped animals only using their phenotypic data (using models (1) as well as (2) and (3) including their sub-models) and for the genotyped animals using their pre-corrected phenotypic data obtained based on adjustment factors estimated from the entire data set (using models (1) as well as (2) and (3) including their sub-models).

6.2.3 Marker effect distribution

The QTL distribution of the trait influences the efficiency of genomic selection. To understand the underlying QTL distribution, a Bayesian approach was used to investigate whether markers with sufficient large effect were present for these traits. This Bayesian approach is implemented in the programme iBay (JANSS 2008) and calculates the change in odds from prior to posterior probability (PPOR) for each marker using the following equation:

$$PPOR = (\hat{p}_1 /(1- \hat{p}_1))/(\pi_1/ \pi_0),$$

where $\hat{p}_1$ is the estimate for the posterior probability of the marker to have an effect, π_0 the proportion of markers associated with no effect and π_1 the proportion of markers associated with an effect. Results were plotted per trait for all markers, whereby a PPOR >3.2 can be interpreted as substantial, a PPOR >10 as strong, and a PPOR >100 as decisive evidence for the marker to be associated with an effect (JANSS 2008).

6.2.4 Efficiency of selection

Variance components for polygenic (σ^2_u), genomic (σ^2_a), pen at mixing (σ^2_{pen}), residual (σ^2_e) and total phenotypic (σ^2_p) effects were estimated using information from all animals as well as for a sample of the data that comprised of both genomic and phenotypic information. The variance due to genomic effects is calculated as the sum of the contributions to the genetic variance from each marker, plus all possible covariances due to linkage disequilibrium, taking into account the allele frequencies.

The software iBay required that animals with only phenotypic data had to be excluded from the analysis.

Predictive ability (PA) was determined by cross validation and calculated as the Pearson's correlation between predicted observation and the corresponding realised observation. Realised observation was calculated as the phenotype adjusted for its corresponding estimates of fixed effects and covariates, while the predicted observation was the estimated breeding value, similar to the approach suggested by LEGARRA *et al.* (2008). In addition, the mean square difference (MSD) between the predicted and realised observation was calculated. In a third approach, the PA of the phenotype was calculated as the Pearson's correlation between the observed phenotype and the predicted phenotype based on the estimates of breeding value, fixed effects and covariates.

In the cross validation approach, the data were split into a training and validation data set. The training data set was used to estimate the parameters of the model while the validation data set contained the animals for which the predicted observations were obtained based on the estimated parameters of the training data set. Size of the training set is of importance for the estimation of accurate genetic parameters (GODDARD and HAYES 2009). To ensure a larger size of training data set, the validation data set was limited to 150 animals. A random split of the data into training and validation data set, considering the family structure as described below, was repeated to create ten validation data sets. Each validation data set had a corresponding training data set, which contained the remaining animals.

Two different routines for splitting the data were used: selection within family and selection between families. For selection within family, full sib families were randomly split between training and validation data set so that the validation data set contained one animal from full sib families with three animals, and two animals from full sib families with four or more animals. This approach ensures that the training data set always contained at least two animals from a full sib family. For selection

between families, the data were randomly split keeping full sib families together so that no full sib family would have animals in both data sets simultaneously. As a result, for selection between families there was less close genetic connectedness between training and validation data than within family selection with full sibs in both data sets.

The approximate change in accuracy of model (2) compared to model (1) was estimated using the equation derived by LEGARRA *et al.* (2008). The basic equation to estimate differences in accuracies between models can be described as follows:

$$\Delta r\,(g, \hat{g}) = \Delta r\,(y, \hat{y})/H\Omega,$$

where g and $\hat{g}$ are the total genetic value of the animal and its estimate, respectively, y and $\hat{y}$ are the realised and predicted observation of the animal, and H^2 and Ω^2 are calculated as σ^2_g/σ^2_p and $\sigma^2_{\hat{g}}/(\sigma^2_{\hat{g}}+\sigma^2_{\hat{c}})$, respectively, with estimates for these variance components based on model (1). As described for the predictive ability, this was calculated for selection within as well as selection between families.

6.3 Results

6.3.1 Estimation of SNP effects

Figures 6.1 and 6.2 show the distribution of the change in odds from prior to posterior probability (PPOR) of each SNP based on model (2) using the 2.5% mixture distribution. Because the PPOR based on all models using different mixtures showed similar distributions, with no evidence for SNP effects (PPOR < 3.2) on lesion scores at mixing and post mixing as well as behavioural traits, only the results of one model are shown in here, and only for the traits LSA1, LSC2, RA and RNRA.

(a) LSA1

(b) LSC2

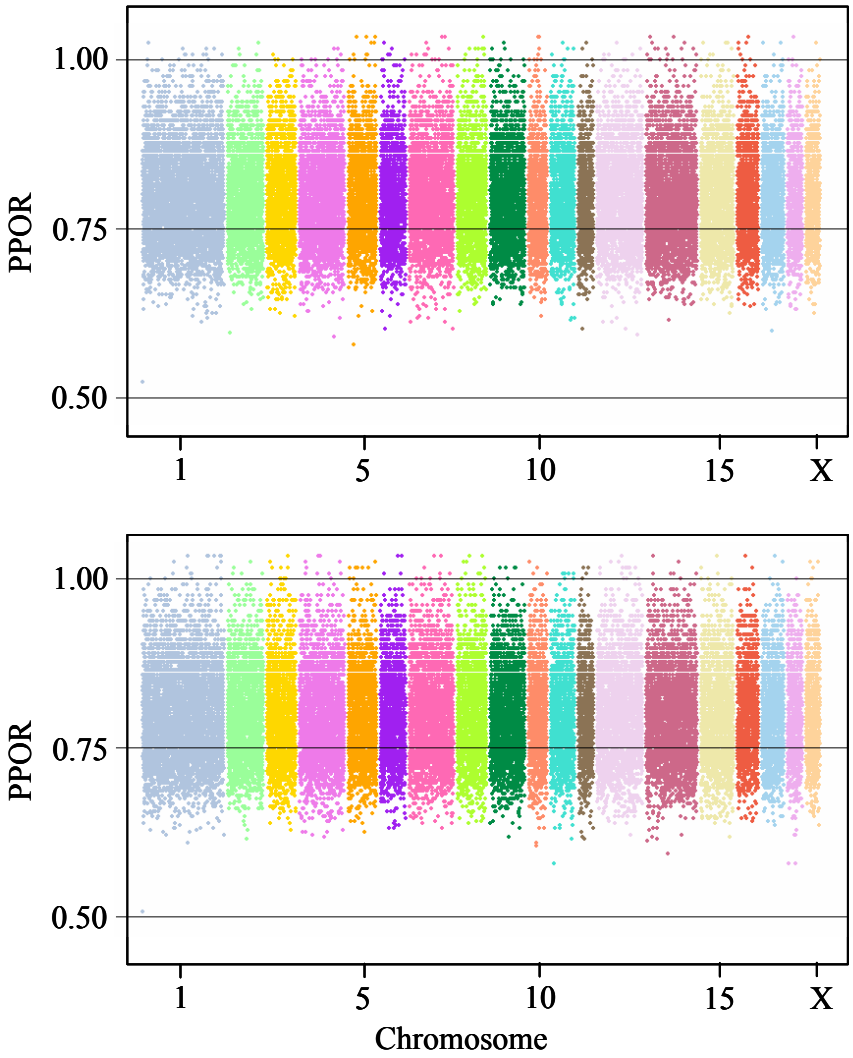


Figure 6.1: Distribution of odds ratios of SNP effects for lesion scores. SNP effects are estimated by the changes in odds from prior to posterior probability (PPOR) using model (2) considering a 2.5% mixture distribution for anterior lesion score at mixing (LSA1) and central lesion score post mixing (LSC2). PPOR < 3.2 denotes not sufficient evidence for a marker to have a significant effect.

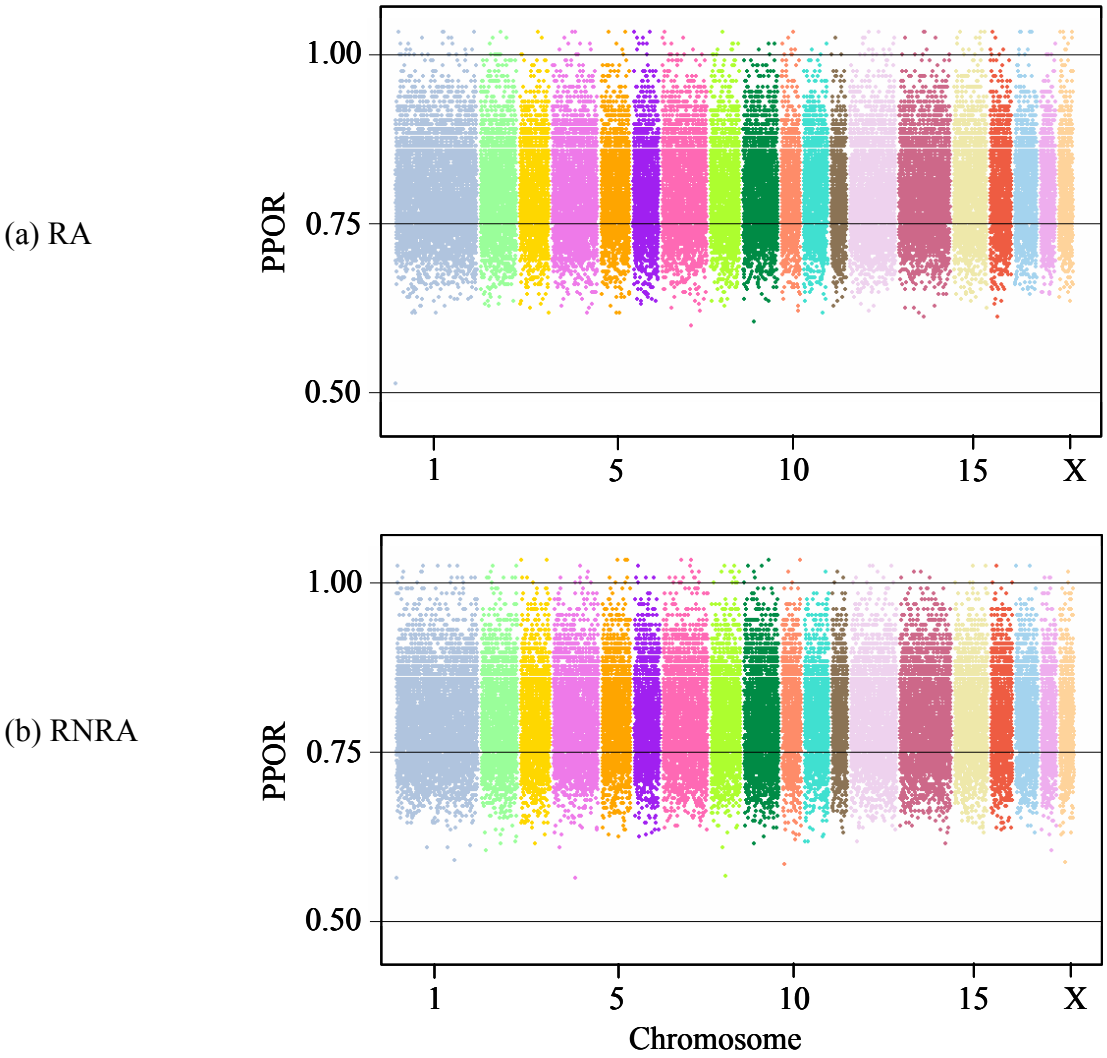


Figure 6.2: Distribution of odds ratios of SNP effects for behavioural traits. SNP effects are estimated by the changes in odds from prior to posterior probability (PPOR) using model (2) considering a 2.5% mixture distribution for reciprocal aggression (RA) and receipt of non-reciprocal aggression (RNRA). PPOR < 3.2 denotes not sufficient evidence for a marker to have a significant effect.

6.3.2 Variance components

Tables 6.2, 6.3 and 6.4 present the estimates for the total phenotypic variances, the heritabilities based on the polygenic effects model, the proportions of variance attributed to genomic effects and the phenotypic proportions of the pen variances. Estimated variance components are based on the entire data set and are presented for six different analyses based on model 1 using the entire data set or a sample data set

including the genotyped animals only, as well as models (2) and (3) considering 100% or 10% of the markers to be associated with an effect on the analysed traits. Results based on sub-models using mixtures of 25%, 2.5% and 1% are not presented, because they showed the same trend.

Table 6.2: Variance components for lesion scores at mixing. Estimates of the total phenotypic variance (σ^2_p), variance attributed to the pen effect (σ^2_{pen}), residual variance (σ^2_e), heritability based on polygenic effects (h^2_u) and proportion of variance attributed to genomic effects (h^2_a) for lesion scores at mixing using the entire data sets or a subset of genotyped animals based on model (1) (2) and (3) (95%-HPD) as well as considering no (100%) and 10% mixture distribution.

Trait	Model/ Data	σ^2_p	σ^2_{pen}	σ^2_e	h^2_u	h^2_a
LSA1	(1) entire	12,080 11,156-13,117	944 447-1434	8264 7145-9354	0.24 0.14-0.34	-
	(1) subset	12,303 10,610-14,235	937 111-1949	8777 6568-10,913	0.21 0.03-0.41	-
	(2) 100%	11,789 10,287-13,395	941 99-1912	9157 7917-10,416	-	0.14 0.11-0.18
	(2) 10%	12,002 10,412-13,711	963 118-1990	8772 7148-10,523	-	0.19 0.03-0.29
	(3) 100%	12,128 10,456-13,922	946 87-1947	8635 6560-10,519	0.11 0.00-0.30	0.10 0.07-0.12
	(3) 10%	12,486 10,688-14,401	956 53-1987	7865 5718-9792	0.11 0.00-0.28	0.18 0.02-0.31
LSC1	(1) entire	12,672 11,533-13,860	1574 884-2342	7907 6613-9114	0.25 0.14-0.37	-
	(1) subset	12,740 10,828-14,847	1511 264-3033	9263 7191-11,357	0.15 0.00-0.33	-
	(2) 100%	12,374 10,569-14,365	1585 370-3148	9541 8225-10,772	-	0.10 0.08-0.13
	(2) 10%	12,423 10,590-14,297	1545 252-2928	9411 7463-11,118	-	0.12 0.01-0.24
	(3) 100%	12,544 10,665-14,567	1548 296-3032	9215 7372-10,919	0.07 0.00-0.21	0.07 0.05-0.08
	(3) 10%	12,670 10,792-14,697	1515 262-2987	9237 7203-11,103	0.12 0.01-0.28	0.03 0.00-0.07
LSE1	(1) entire	9174 8372-9986	1239 745-1802	6232 5475-7034	0.18 0.08-0.28	-
	(1) subset	8892 7314-10,608	1762 508-3201	5648 4414-6815	0.17 0.04-0.31	-
	(2) 100%	8801 7253-10,412	1835 659-3364	5838 5063-6685	-	0.13 0.10-0.16
	(2) 10%	8731 7252-10,436	1821 587-3316	5988 5036-6968	-	0.11 0.02-0.18
	(3) 100%	9000 7450-10,761	1846 629-3396	5442 4178-6600	0.09 0.00-0.23	0.10 0.08-0.13
	(3) 10%	8916 7369-10,700	1819 597-3312	5544 4372-6653	0.12 0.01-0.26	0.05 0.01-0.12

Analysis based on model (1) using the entire data set showed heritabilities that were similar to those found by TURNER *et al.* (2009) (Tables 6.1 and 6.2). Generally, the lesion scores in the caudal region resulted in lowest heritability at mixing and post mixing, while the anterior and central region had higher heritabilities. Of the behavioural traits, RA resulted in the highest heritability and RNRA gave the lowest. Using model (1) and the genotyped data only, the heritabilities were not different from those using the entire data for most traits, except for LSA2 where the heritabilities based on those data sets were 0.14 (0.01-0.39) and 0.43 (0.30-0.57), respectively. Moreover, the trait LSC1 showed a substantially lower heritability of 0.15 (0.00-0.0.33) compared to 0.25 (0.14-0.37) for the same data sets.

Phenotypic proportions of the pen effect ranged from 4% to 14% for all nine traits using model (1) and the entire data set (Tables 6.2 to 6.4). Using model (1) and the subset of genotyped animals only, the phenotypic proportion of the pen effect increased slightly for most traits, except for LSA1 and LSC1 which did not change and for LSE2 which decreased. Proportions of the variance attributed to the genomic effect based on model (2) considering all SNPs influencing the trait (100%) were 33% to 45% lower for the lesion scores at mixing and post-mixing in the anterior and central region, 24% to 26% lower for corresponding lesion scores in the caudal region and 30% to 38% lower for the three behavioural traits, compared to model (1) using the same subset of data (Tables 6.2 to 6.4). For all traits this change in proportion was compensated by an increase in the phenotypic proportion of the residual effect of 3% to 18%-points and a slight increase in the phenotypic proportion of the pen effect of up to 2%-points.

Using mixture distribution in model (2), considering only 10% of the SNPs affecting the traits, resulted in an decrease in proportion of the variance attributed to the genomic effect compared with model (2) considering all SNPs affecting the traits (100%) for most traits, except for LSA1, LSC1 and RNRA. However, with decreasing mixtures percentages, 95%-HPD intervals increased rapidly. Differences among the four mixture models (1%, 2.5%, 10% and 25%, results not shown) were

small and mainly not significant. For example, the three lesion scores at mixing showed proportion of the variance attributed to the genomic effect for the 1% mixture distribution that were intermediate between non-mixture (100%) and the 10% mixture, but with substantial larger 95%-HPD intervals (results not shown).

Table 6.3: Variance components for lesion scores post mixing. Estimates of the total phenotypic variance (σ^2_p), variance attributed to the pen effect (σ^2_{pen}), residual variance (σ^2_e), heritability based on polygenic effects (h^2_u) and proportion of variance attributed to genomic effects (h^2_a) for lesion scores at three weeks post mixing using the entire data sets or a subset of genotyped animals based on model (1) (2) and (3) (95%-HPD) as well as considering no (100%) and 10% mixture distribution.

Trait	Model/ Data	σ^2_p	σ^2_{pen}	σ^2_e	h^2_u	h^2_a
	(1) entire	3149 2852-3459	267 139-418	1519 1186-1835	0.43 0.30-0.57	-
	(1) subset	2604 2193-3037	378 86-733	1866 1478-2255	0.14 0.01-0.29	-
LSA2	(2) 100%	2537 2148-2954	385 96-734	1952 1698-2222	-	0.08 0.06-0.10
	(2) 10%	2523 2131-2950	389 105-746	2087 1818-2373	-	0.02 0.00-0.06
	(3) 100%	2610 2188-3027	376 101-718	1798 1373-2147	0.11 0.00-0.28	0.05 0.04-0.07
	(3) 10%	2637 2243-3112	374 93-716	1712 1292-2107	0.09 0.00-0.23	0.12 0.03-0.24
	(1) entire	3465 3159-3783	363 203-546	1937 1595-2234	0.33 0.22-0.45	-
	(1) subset	3229 2658-3849	405 64-794	1447 801-2107	0.42 0.16-0.68	-
LSC2	(2) 100%	3001 2565-3476	418 98-822	1885 1589-2164	-	0.23 0.18-0.28
	(2) 10%	2936 2510-3443	431 101-836	2099 1737-2479	-	0.14 0.03-0.23
	(3) 100%	3266 2701-3889	403 65-789	1375 722-2051	0.32 0.04-0.57	0.13 0.10-0.16
	(3) 10%	3237 2651-3882	414 86-833	1428 719-2114	0.40 0.13-0.68	0.02 0-0.08
	(1) entire	4513 4172-4896	418 221-634	3345 2937-3748	0.17 0.07-0.26	-
	(1) subset	4572 3957-5231	178 0-459	3514 2710-4372	0.19 0.02-0.39	-
LSE2	(2) 100%	4460 3914-5035	182 1-465	3676 3191-4209	-	0.14 0.11-0.17
	(2) 10%	4385 3839-4945	176 0-449	4098 3487-4695	-	0.03 0.00-0.13
	(3) 100%	4602 3983-5296	182 0-460	3353 2496-4081	0.13 0.01-0.33	0.10 0.07-0.12
	(3) 10%	4563 3919-5211	178 0-464	3519 2675-4357	0.16 0.00-0.35	0.02 0.01-0.04

Table 6.4: Variance components for behavioural traits. Estimates of the total phenotypic variance (σ^2_p), variance attributed to the pen effect (σ^2_{pen}), residual variance (σ^2_e), heritability based on polygenic effects (h^2_u) and proportion of variance attributed to genomic effects (h^2_a) for behavioural traits using the entire data sets or a subset of genotyped animals based on model (1) (2) and (3) (95%-HPD) as well as considering no (100%) and 10% mixture distribution.

Trait	Model/ Data	σ^2_p	σ^2_{pen}	σ^2_e	h^2_u	h^2_a
RA	(1) entire	41,192 36,978-45,885	2048 450-3783	21,543 16,002-26,542	0.42 0.27-0.58	-
	(1) subset	39,173 31,902-46,754	5907 1485-11,484	18,984 11,945-25,907	0.36 0.15-0.58	-
	(2) 100%	37,193 31,563-43,818	6187 1602-11,975	22,540 19,194-26,061	-	0.23 0.18-0.28
	(2) 10%	36,993 31,417-44,116	6073 1369-11,537	22,934 18,106-28,413	-	0.22 0.04-0.33
	(3) 100%	39,240 32,622-46,788	6217 1513-12,106	17,928 11,306-24,143	0.24 0.04-0.44	0.14 0.10-0.18
	(3) 10%	39,256 32,379-47,077	5991 1359-11,587	18,210 11,124-25,475	0.31 0.11-0.56	0.07 0.00-0.19
DNRA	(1) entire	37,154 33,271-40,872	2683 966-4628	22,394 17,975-26,811	0.32 0.20-0.47	-
	(1) subset	39,101 32,893-46,167	4105 453-8693	23,193 16,095-29,822	0.30 0.09-0.5	-
	(2) 100%	37,410 32,617-43,115	4059 581-8400	25,415 21,847-29,415	-	0.21 0.17-0.26
	(2) 10%	37,007 31,595-42,933	3755 353-7881	27,251 20,346-34,966	-	0.16 0.01-0.34
	(3) 100%	39,554 33,626-46,629	4255 623-8862	21,539 15,235-27,688	0.20 0.03-0.38	0.14 0.11-0.18
	(3) 10%	39,253 33,160-46,300	4115 458-8499	22,101 15,239-29,426	0.21 0.01-0.42	0.12 0.02-0.24
RNRA	(1) entire	21,730 19,531-23,897	3138 1652-4742	16,771 14,788-18,633	0.08 0.02-0.15	-
	(1) subset	22,990 19,605-27,230	3588 939-6588	16,422 13,086-19,654	0.13 0.01-0.27	-
	(2) 100%	22,529 19,248-26,441	3592 1081-6653	17,078 14,899-19,398	-	0.08 0.06-0.10
	(2) 10%	22,693 19,357-26,657	3623 1102-6704	17,042 14,336-19,547	-	0.09 0.02-0.17
	(3) 100%	23,011 19,603-27,104	3568 1066-6736	16,092 12,979-19,293	0.09 0.00-0.23	0.05 0.04-0.06
	(3) 10%	23,010 19,446-27,144	3594 968-6742	16,164 12,797-19,115	0.10 0.00-0.22	0.04 0.01-0.09

Using a model (3) which considered both the polygenic and genomic effects, the polygenic component effectively captured the part of the genetic variance that was not accounted for by the genomic component so that the total variance attributed to genetic effects (genomic and polygenic combined) was similar to the polygenic

variance estimated by using model (1). Proportions of the variance attributed to the genomic effect using model (3) were slightly lower than using model (2) for all traits. They were 52% to 69% lower for the lesion scores at mixing and post-mixing in the anterior and central region, 41% to 47% lower for the lesion scores in the caudal region and 53% to 62% lower for the behavioural traits, compared to model (1) using the corresponding data set of genotyped animals only. Using model (3), the heritabilities due to the polygenic effects accounted for 47% to 53%, 68% to 79%, and 67% to 69% of the heritability of the lesion score traits at mixing, at post mixing, and the behavioural traits, respectively, estimated using model (1).

Using a mixture distribution in model (3) lead to lower proportions of the variance attributed to the genomic effect for many traits, but estimates varied among mixtures and, in contrast to the results presented in Chapter 4, no clear trend for a decrease with decreasing percentages of SNPs considered to affect the traits was visible.

6.3.3 Predictive ability

In the following, description of the results will focus on four of the nine traits, namely LSA1, LSC2, RA and RNRA. Results of LSC1 and LSE1 showed similar patterns to those of LSA1 but with slightly different values. Results of LSA2 and LSE2 showed similar patterns to those of LSC2 and DNRA similar patterns to those of RA but with slightly lower values. These five traits will therefore only be discussed when clearly different from the other four traits.

Table 6.5 presents the average PA of the phenotype and genotype using model (1) for selection within family (W) or between families (B) for the traits LSA1, LSC2, RA and RNRA. The results indicate that selection within family always outperformed selection between families for all traits. The PA of the phenotype was always higher than the PA of the genotype and showed less variation among traits. The PA increased with increasing heritability, so that the lowest heritable trait RNRA showed the lowest PA, while the highest heritable trait RA resulted in the highest PA.

Table 6.5: Predictive ability of the phenotype and genotype using the entire data set. Average predictive ability using the entire (E) data set for model (1) for selection within (W) or between (B) families for lesion score and behavioural traits.

	LSA1		LSC2		RA		RNRA	
Model/Criteria	W [a]	B [a]	W [b]	B [b]	W [a]	B [b]	W [b]	B [b]
(1) phenotype^{E}	0.29	0.12	0.37	0.22	0.39	0.19	0.24	0.19
(1) genotype^{E}	0.17	0.02	0.28	0.06	0.33	0.14	0.09	0.01

[a] s.e. ≤ 0.02; [b] s.e. ≤ 0.03

Table 6.6 summarises the average PA of the genotype, using 552 genotyped animals only. Again, selection within family outperformed selection between families, except for the trait RNRA. All lesion scores showed an intermediate predictive ability when selection was within family. All traits showed similar PA for models (2) and (3) compared to model (1), with generally a slight but non-significant increase in PA. LSC2 was the only exception, showing a decrease in PA for model (2) compared to model (1). However, this was not obtained when comparing model (3) to model (1). When selection was carried out between families, a similar picture was visible, with models (2) and (3) showing slightly higher PA than model (1) except for LSC2. However, this increase was only significant for model (2) for LSA1, LSC1 and RNRA.

Comparing the use of different mixture distribution in the models, no trait showed a significant decrease in PA with decreasing percentages when selection was within family. Only LSA1 and LSC2 showed a slight trend for a decrease in PA for selection within family for mixtures below 10% in model (2). RA and LSE1 (not shown) showed a trend for a decrease in PA in model (3) for mixtures below 25%. For selection between families, LSA1 was the only trait to show a tendency for a decrease in PA using model (3).

Table 6.6: Predictive ability of the genotype using the subset of genotyped animals only. Average predictive ability of all tested models and mixture distributions using 552 genotyped animals for selection within (W) or between (B) families for lesion score and behavioural traits.

	LSA1		LSC2		RA		RNRA	
Model/Mixture	W [b]	B [c]	W [b]	B [b]	W [a]	B [c]	W [b]	B [c]
(1)	0.15	0.06	0.26	0.04	0.27	0.11	0.05	0.00
(2) 100%	0.20	0.16	0.22	-0.01	0.30	0.13	0.04	0.07
(2) 25%	0.20	0.16	0.20	-0.01	0.29	0.13	0.04	0.08
(2) 10%	0.20	0.15	0.19	-0.01	0.28	0.13	0.04	0.08
(2) 2.5%	0.18	0.15	0.20	-0.01	0.28	0.13	0.04	0.08
(2) 1%	0.17	0.16	0.19	-0.01	0.29	0.13	0.04	0.08
(3) 100%	0.20	0.13	0.26	0.02	0.31	0.14	0.06	0.04
(3) 25%	0.20	0.12	0.26	0.02	0.30	0.14	0.05	0.03
(3) 10%	0.20	0.12	0.26	0.01	0.29	0.13	0.05	0.04
(3) 2.5%	0.20	0.12	0.25	0.03	0.30	0.13	0.06	0.04
(3) 1%	0.19	0.09	0.26	0.03	0.29	0.13	0.06	0.03

[a] s.e. ≤ 0.01; [b] s.e. ≤ 0.02; [c] s.e. ≤ 0.03

Table 6.7 gives the average PA of the genotype, using 552 animals and pre-adjusted phenotypic data for the fixed effects and covariables. The difference to other analyses is that the adjustment factors were predicted on the entire data set. Pre-adjustment of the data generated largely the same results, with slightly higher PA for most traits except RNRA compared to using the original phenotypic data, but these differences were not significant.

Table 6.7: Predictive ability of the genotype using pre-adjusted observations of the subset of genotyped animals only. Average predictive ability of all tested models and mixture distributions using 552 genotyped animals for selection within (W) or between (B) families for lesion score and behavioural traits.

	LSA1		LSC2		RA		RNRA	
Model/Mixture	W [a]	B [a]	W [a]	B [b]	W [a]	B [a]	W [a]	B [b]
(1)	0.17	0.07	0.28	0.10	0.30	0.18	0.04	-0.01
(2) 100%	0.22	0.17	0.24	0.01	0.32	0.18	0.03	0.06
(2) 25%	0.23	0.16	0.22	0.01	0.31	0.17	0.03	0.06
(2) 10%	0.21	0.16	0.21	0.01	0.30	0.18	0.03	0.06
(2) 2.5%	0.22	0.17	0.20	0.00	0.30	0.17	0.03	0.06
(2) 1%	0.21	0.17	0.22	0.01	0.30	0.17	0.03	0.07
(3) 100%	0.21	0.13	0.28	0.06	0.33	0.20	0.04	0.03
(3) 25%	0.21	0.13	0.28	0.08	0.32	0.19	0.04	0.03
(3) 10%	0.20	0.13	0.28	0.08	0.32	0.19	0.04	0.03
(3) 2.5%	0.19	0.13	0.28	0.08	0.31	0.19	0.04	0.01
(3) 1%	0.20	0.11	0.28	0.07	0.31	0.20	0.04	0.02

[a] s.e. ≤ 0.02; [b] s.e. ≤ 0.03

Table 6.8 presents the PA of the phenotype using all 552 genotyped animals. Similar to Table 6.5, PA of the phenotype was always higher than PA of the genotype. Interestingly, while the PA of the genotype was generally stable across mixtures, the PA of the phenotype showed a trend for a decrease across mixtures for several traits. LSA1 displayed a trend for a decrease in PA for selection within family for mixtures below 10% in model (2). For RA and DNRA (not shown), a decrease for selection within family for mixtures below 25% was visible using model (2), but this was followed by a slight increase again at 1%. When selection was between families, no change in PA of the phenotype with decrease in mixture percentage was found.

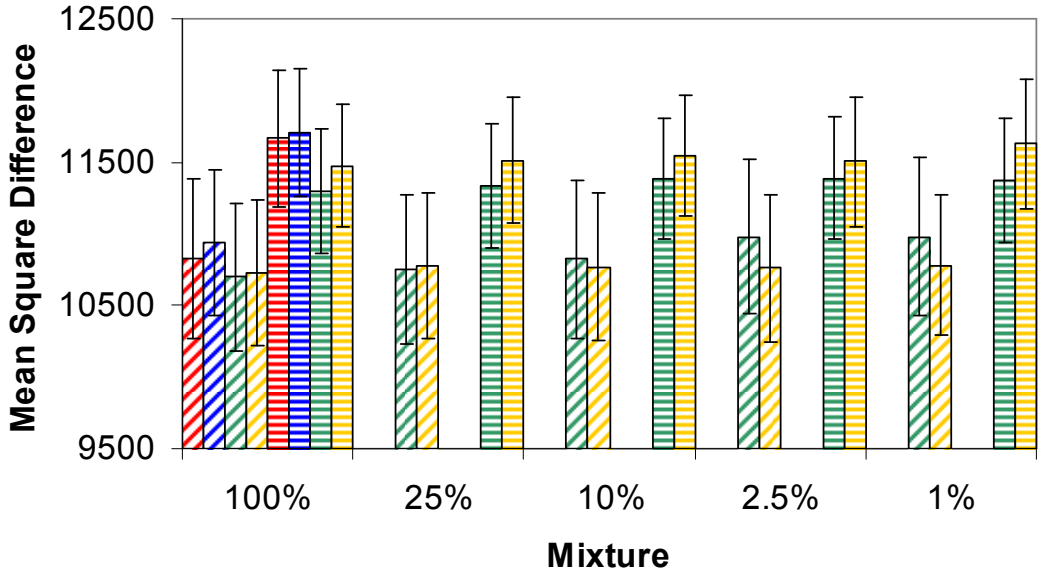
Table 6.8: Predictive ability of the phenotype using a subset of genotyped animals only. Average predictive ability of all tested models and mixture distributions using 552 genotyped animals for selection within (W) or between (B) families for lesion score and behavioural traits.

Trait	LSA1		LSC2		RA		RNRA	
Model/Mixture	W [a]	B [c]	W [b]	B [b]	W [a]	B [a]	W [b]	B [a]
(1)	0.27	0.17	0.36	0.21	0.38	0.24	0.19	0.15
(2) 100%	0.30	0.21	0.34	0.17	0.40	0.26	0.20	0.16
(2) 25%	0.30	0.20	0.32	0.18	0.38	0.24	0.19	0.15
(2) 10%	0.29	0.19	0.32	0.18	0.36	0.24	0.19	0.16
(2) 2.5%	0.26	0.19	0.33	0.17	0.36	0.25	0.20	0.16
(2) 1%	0.26	0.20	0.32	0.18	0.38	0.25	0.20	0.16
(3) 100%	0.30	0.19	0.36	0.19	0.41	0.26	0.20	0.15
(3) 25%	0.30	0.19	0.36	0.19	0.40	0.26	0.19	0.15
(3) 10%	0.30	0.19	0.36	0.19	0.40	0.25	0.20	0.15
(3) 2.5%	0.30	0.19	0.36	0.20	0.40	0.25	0.20	0.15
(3) 1%	0.29	0.18	0.36	0.20	0.39	0.25	0.20	0.15

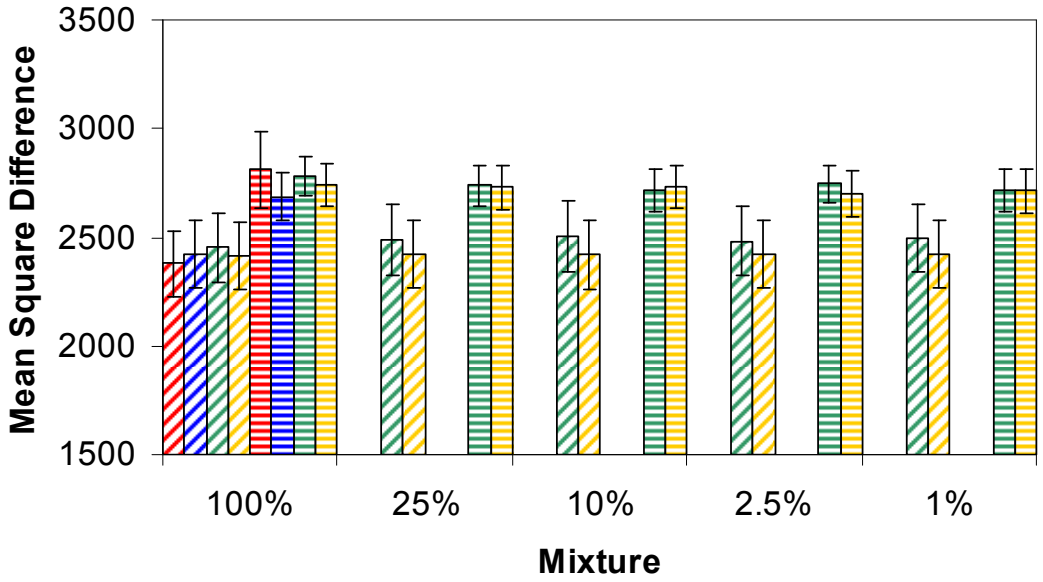
[a] s.e. ≤ 0.02; [b] s.e. ≤ 0.03; [c] s.e. ≤ 0.04

6.3.4 Mean square difference

Figures 6.3 and 6.4 show the MSD between the predicted and realised observations for the traits LSA1, LSC2, RA and RNRA. In contrast to the previous paragraph, where a higher value of PA is favourable, a lower MSD is an indicator for better performance of a model. The models are comparable to model (1) using the "genotype^{E}" results (Table 6.5) and all models in Table 6.6, using a subset of data with genotyped animals only.



(a) LSA1



(b) LSC2

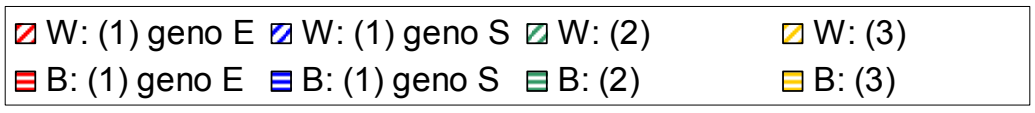
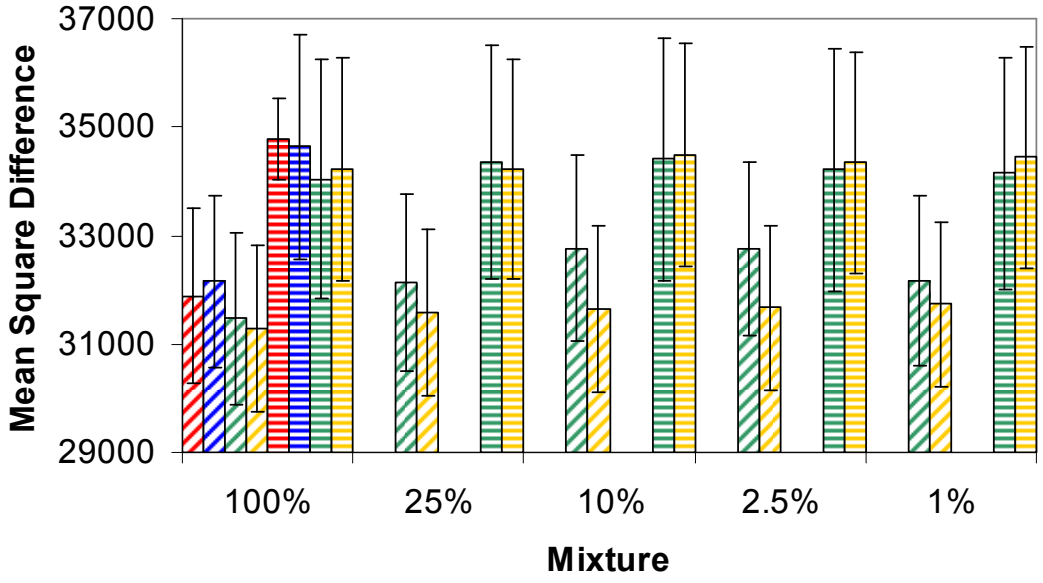
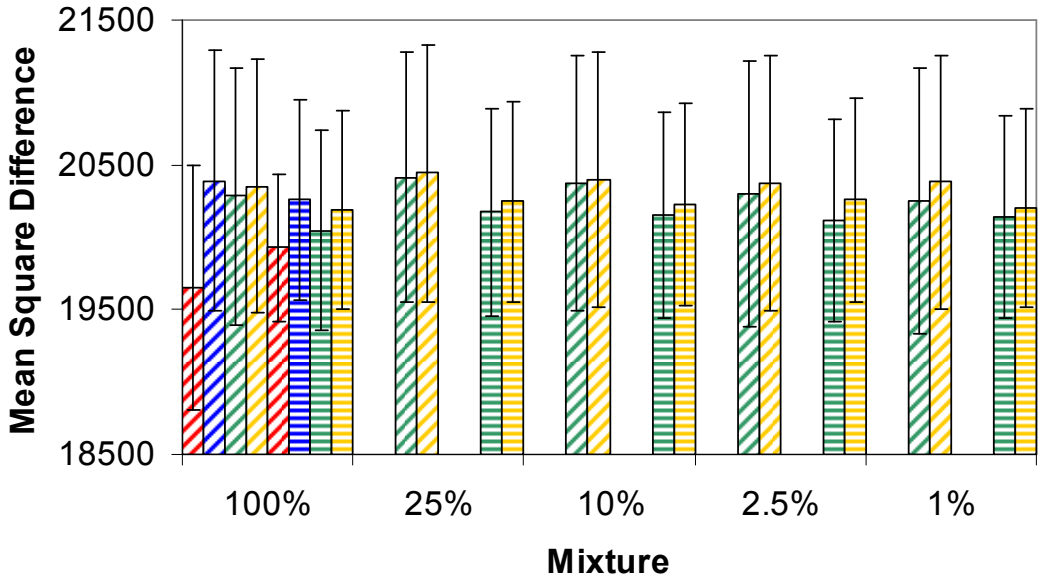


Figure 6.3: Mean square differences between predicted and realised observations for lesion scores. Mean square differences for model (1) using the entire data – "(1) geno E" – and model 1 to 3 using a subset of data of genotyped animals only –" (1) geno S", "(2)" and "(3)" – for selection within (W) or between (B) families for (a) anterior lesion score at mixing (LSA1) and (b) central lesion score post mixing (LSC2).



(a) RA



(b) RNRA

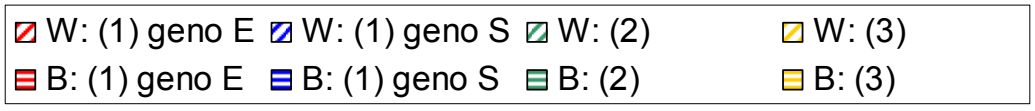


Figure 6.4: Mean square differences between predicted and realised observations for behavioural traits. Mean square differences for model (1) using the entire data – "(1) geno E" – and model 1 to 3 using a subset of data of genotyped animals only – "(1) geno S", "(2)" and "(3)" – for selection within (W) or between (B) families for (a) reciprocal aggression (RA) and (b) receipt of non-reciprocal aggression (RNRA).

In agreement with the results of PA in the previous paragraph, selection within family (diagonally patterned bars) performed better than selection between families (horizontally patterned bars) except for models (2) and (3) for RNRA. However, RNRA showed the least differences in MSD between selection within and between families. For LSA1 and RA, models that included genomic information showed a lower MSD than models excluding this information when no mixture distribution was fitted. A decrease in mixture percentage for these traits affected the MSD for model (2), but did not change the MSD for model (3). Both models (1) genotype^{E} by and large showed a slightly lower MSD and thus a better estimate than models (1) genotype^{S}.

6.3.5 Accuracy

Table 6.9 presents the approximate increase in accuracy for model (2) compared to model (1) for LSA1, LSC2, RA and RNRA for all mixtures. For selection within family, a gain in accuracy across all mixture percentages was found for most traits, ranging from 0.00 in LSA2 to 0.16 in LSC2 (not shown). LSC2, LSE2 (not shown) and RNRA showed decreases in accuracies across all mixture percentages, ranging up to -0.12 in LSC2. Including genomic information was of greater benefit (or smaller disadvantage) when selection was between families, compared to selection within family, for all traits except LSC1 and RA. Across mixtures, LSA1 and LSC1 showed a tendency for a decreased gain in accuracy with lower mixture percentages when selection was within family. RNRA showed a tendency for an optimum percentage at 2.5% for between family selection where the gain in accuracy was the highest.

Table 6.9: Estimated differences in accuracies for the genotype using all genotyped animals. Estimated increase of accuracy of model (2) compared to model (1) as used in Table 6.6, using 552 genotyped animals for selection within (W) or between (B) families for lesion score and behavioural traits.

	LSA1		LSC2		RA		RNRA	
Model/Mixture	W[c]	B[e]	W[a]	B[d]	W[b]	B[c]	W[d]	B[f]
(2) 100%	0.13	0.24	-0.06	-0.07	0.05	0.03	-0.05	0.34
(2) 25%	0.14	0.25	-0.10	-0.08	0.04	0.03	-0.06	0.36
(2) 10%	0.12	0.23	-0.12	-0.09	0.01	0.04	-0.06	0.38
(2) 2.5%	0.06	0.22	-0.10	-0.08	0.02	0.04	-0.08	0.39
(2) 1%	0.06	0.25	-0.11	-0.09	0.03	0.02	-0.07	0.35

[a] s.e. ≤ 0.02; [b] s.e. ≤ 0.03; [c] s.e. ≤ 0.04; [d] s.e. ≤ 0.05; [e] s.e. ≤ 0.06; [f] s.e. ≤ 0.09;

6.4 Discussion

6.4.1 SNP effects

Research on aggressive behaviour in different species has found few QTL. Substantial research has been done in animals that serve as model species such as Drosophila (five QTL - EDWARDS and MACKAY 2009), mice (two QTL - BRODKIN *et al.* 2002) or rats (two QTL - ALBERT *et al.* 2009). The QTL in the last study explained 2.3% and 5.1% of the residual phenotypic variance. The latter two studies included animals from crosses between lines that had been selected for aggressive behaviour for many generations. Analysing dogs, VAN DEN BERG *et al.* (2008) found no markers associated with either dog-directed or human-directed aggressiveness.

In livestock species, a form of aggressive behaviour that has been studied extensively is feather pecking. BUITENHUIS *et al.* analysed records from 630 laying hens and detected three suggestive and one significant QTL for delivering feather pecking (BUITENHUIS *et al.* 2003b), five suggestive and one significant QTL for receiving feather pecking and three suggestive QTL for receiving toe picking (BUITENHUIS *et al.* 2003a). Moreover, JENSEN *et al.* (2005) identified one suggestive QTL for feather

pecking, but the residual variance explained by this QTL was low at 0.7%. A recent study of ten candidate genes for feather pecking in two lines of laying hens found four genes which were differentially expressed (WYSOCKI *et al.* 2010). For maternal infanticide in pigs, another form of aggressive behaviour, four significant QTL were reported by (CHEN *et al.* 2009).

All these studies on traits related to aggressive behaviour reported only a few significant QTL and these QTL explained only a small percentage of the phenotypic variance. None of the SNP markers analysed showed sufficient evidence to affect the aggressive behaviour analysed in the present data. The effect of a QTL may be spread over several markers in a region, whereby each individual marker picks up part of the effect of the QTL. Moreover, analyses of the present data set using four different approaches did not find any significant QTL (R. PONG-WONG, The Roslin Institute and R(D)SVS, The University of Edinburgh, United Kingdom, personal communication). However, analysis at chromosomal level, using 18 genetic variance components, found that a few chromosomes harbour the genetic variation for these traits (R. PONG-WONG, unpublished results). TERENINA *et al.* (2010) genotyped animals for aggressive behaviour using the standardised approach validated by TURNER *et al.* (2006b) for candidate genes that are known to be associated with aggressive behaviour. They detected new polymorphisms in the candidate genes and further association studies are in progress.

The study on mice indicated that the methodology used in the present analysis was able to find differences between the traits in the numbers of markers with an effect. These differences correspond with the relative differences in number of QTL in these traits (Chapter 4, Table 4.9). Moreover, the number of animals genotyped may be not large enough to identify significant QTL in the present data set.

6.4.2 Heritabilities

Estimates of heritabilities for all traits were similar to those reported by TURNER *et al.* (2009), using the same data set but different methodology. Whereas TURNER *et al.*

(2009) used the Frequentist's approach based on REML, the present study used a Bayesian approach, which may explain the slight difference between estimates. Reducing the data set to 552 animals did not change these heritabilities for most traits, except for LSC1 and LSA2. Differences between the analysis of the entire data set and the reduced data set may be due to the fact that cross bred animals were included in the entire data set, whereas the reduced data set only contained purebred animals. The model did not consider dominance and heterosis, which may have affected the estimates if present. The heritabilities obtained are slightly higher than those found for other welfare-related traits, e.g. piglet survival, ranging from 0.01 to 0.13 (ROSENDO *et al.* 2007b; SU *et al.* 2007; ROEHE *et al.* 2009), but lower than heritabilities found for some production traits e.g. backfat thickness, ranging from 0.30 to 0.51 (KNOL 2001; SERENIUS and STALDER 2004; ZUMBACH *et al.* 2007).

Heritabilities for behavioural traits are often of low to moderate magnitude. Studies by LØVENDAHL *et al.* (2005), TURNER *et al.* (2008) and D'EATH *et al.* (2009) reported heritabilities ranging from 0.04 to 0.47 for various aggressiveness traits in pigs. Heritabilities for different forms of aggressive behaviour in other species showed a much larger range, from 0.06 in cattle (PHOCAS *et al.* 2006) up to 0.81 in dogs (LIINAMO *et al.* 2007). Studies in species such as fish and Drosophila reported low heritabilities for aggressive behaviour, ranging from 0.01 to 0.14 (BELL 2005; EDWARDS *et al.* 2006). However, divergent selection experiments in a broad range of species, including mice, rats and foxes, have shown that selection for or against aggressive behaviours can be extremely successful (HARE *et al.* 2005; EDWARDS *et al.* 2006; ALBERT *et al.* 2009). Based on the heritabilities estimated in this study, selection against aggressive behaviour, either through selection for behaviour directly or via selection for lesion scores, should be successful.

6.4.3 Efficiency of selection

The second objective of the work described in this chapter was to investigate the efficiency of genomic selection for behavioural traits compared to traditional polygenic selection. Research has shown that lesion scores in the anterior region may

be a good indicator for RA as well as DNRA (TURNER *et al.* 2009). Both RA and DNRA showed a slightly higher PA than the lesion score traits LSA1 and LSA2. However, including a genomic component was of greater benefit for the two lesion scores, especially for LSA1, with equally high or higher increases of accuracy compared to the behavioural traits. The aforementioned study also found that lesion scores in the central or caudal region may be a good indicator for RNRA (TURNER *et al.* 2009). All four lesion score traits showed a higher PA than RNRA, especially when selection was within family, but small or negative gains in accuracy. Inclusion of a genomic component benefited the accuracy of RNRA greatly when selection was between families. Overall, the results for PA and accuracy gains differed greatly between traits. Likewise, R. Pong-Wong compared genomic selection with best linear unbiased prediction (BLUP) and found the increase in accuracy of GS relative to BLUP to be very diverse between traits, with no or negative gains for some traits (R. PONG-WONG, unpublished results).

The low numbers of observations in this study may play an important role in the variable results found between traits. Studies have shown that accuracies of genomic breeding values depend on the size of the reference population (reviewed by GODDARD and HAYES 2009). In addition, aggressive behaviours are complex traits, influenced by many different genes with a relatively small effect, thus requiring more observations to estimate these accurately. Overall, inclusion of genomic selection may improve selection on either behavioural measurements or lesion scores and thus selection against aggressive behaviour. However, the differences between the traits indicate that this should be considered carefully per trait and breeding scheme.

Genomic selection has been applied successfully in dairy cattle breeding (HAYES *et al.* 2009; VANRADEN and SULLIVAN 2010), but to date it has not been implemented widely in other species. The cost of genotyping may play a role in the lower uptake of genomic selection in species other than dairy cattle, where the relative value of an individual animal is high (GODDARD and HAYES 2009). A simulation study into the benefit of genomic selection for disease resistance in pigs found that selection for a

low-heritable trait with moderate economic value would generate small genetic gains when no correlations with economically important traits existed, but the much higher emphasis on the other traits in a pig breeding programme made progress in disease resistance difficult (HENRYON *et al.* 2010).

All traits investigated in this chapter showed predictive abilities that were in range of those for traits with similar heritabilities in mice (Chapter 4). As discussed previously in this chapter, no significant SNP effects were found for the traits. Chapter 4 showed that the QTL distribution had a clear influence in the efficiency of genomic selection. Simulation studies (KIZILKAYA *et al.* 2009; ZHONG *et al.* 2009) have shown that the number of QTL affecting a trait influences the performance of genomic selection, though the influence differed depending on the method that was used. Commonly, with fewer large-effect QTL, the efficiency of genomic selection decreased. Due to the lack of evidence for significant markers found in the present pig data for all nine traits, no difference in efficiency of genomic selection due to variation in SNP effects distribution was visible.

In general, a decrease in heritability led to a decrease in predictive ability for all traits. An exception was LSE2, which showed a lower PA across models than LSC1 and LSE1 despite a slightly higher heritability. Most methodologies for genomic selection are limited to genotyped animals only, and information from animals with phenotypic records only is discarded. Pre-adjustment of the data allowed the use all animals to predict the influence of fixed effects, but did not change the predictive ability significantly. Inclusion of a polygenic effect had no significant influence on the performance of genomic selection, as was found also in studies by LEGARRA *et al.* (2008) and DE LOS CAMPOS *et al.* (2009), as well as in Chapter 4.

For most traits, selection within family performed better than selection between families, due to information from full sibs. For selection between families, the closest relatives in the present data were half sibs. In Chapter 4, no half sibs were available and selection between families benefited considerably when including genomic

selection, especially for traits with lower heritabilities and no QTL. In this chapter, even though information from relatives in the form of half sibs was available, the increased accuracy of model (2) compared to model (1) was still significant for most traits. This analysis, in which half sib information is available, is a better reflection of the actual situation in a pig breeding programme.

Reducing the number of markers allowed to have an effect had little influence on the efficiency of genomic selection for most traits. LSA1 and LSC2 showed a trend for a decrease in PA for selection within family when percentages were below 10% using genomic information only. When both genomic and polygenic information was included, RA and LSE1 showed a trend for a decrease in PA for mixtures below 25%. This suggests that, depending on the trait, a subset of markers may be sufficient to provide a cost-effective strategy for genomic selection. As discussed in the previous section, the relative value of an individual dairy cow may have played an important role in the high uptake of genomic selection in this species. Other species, including in particular pigs, but also species such as chicken, have a much shorter generation interval. As a result, numerous animals have to be genotyped, thus increasing the costs of genotyping considerably. This makes strategies that reduce the number of animals or the density of the SNP panel much more valuable in these breeding programmes.

A study by CLEVELAND *et al.* (2010) on the reproduction traits total number of piglets born and percentage of stillborn found high accuracies for genomic breeding values for these traits. When a selected subset of markers was used, accuracies decreased slightly for total number born, but decreased much more for percentage of still born, suggesting that there is difference in the genetic structure between these two traits. DEEB *et al.* (2010) genotyped pigs for markers chosen for their assumed association with scrotal hernia in pigs and used a subset of these markers for further validation. Using these selected markers, the accuracy of breeding values is expected to increase by 70% to 100% over using phenotypic information only.

6.5 Conclusion

In this study, no SNP markers showed sufficient evidence to have a significant effect on aggressive behaviour in pigs. Heritabilities for all traits were low to moderate, and models that included a genomic effect generally performed better than models using a traditional polygenic effect only. Especially, in situations where no full sib information was available, genomic selection provided considerable increases in accuracy for some traits. Models including a polygenic effect captured more of the genetic variance, but did not improve the predictive ability of the models. The consistent performance of genomic selection across different mixture percentages indicates that lower numbers of markers are still sufficient to predict observations.

Genomic selection for traits related to aggressive behaviour in pigs is expected to be beneficial. Lesion scores at mixing showed higher predictive abilities than receipt of reciprocal aggression, and a higher gain in accuracy than reciprocal aggression and delivery of non-reciprocal aggression. Genomic selection showed a consistently high efficiency across different mixture distributions with decreasing numbers of SNPs. Therefore, depending on the combination of traits, a low-density SNP panel may be sufficient for a multi-trait approach which considers not only aggressive traits but also other economically important traits. A reduced density of the SNP panel would lower the high costs associated with genotyping animals in pig breeding with its short generation interval. Depending on genetic correlations between these traits, genomic selection against aggressive behaviour in pigs may be a viable option to improve animal welfare in pig breeding programmes.

Chapter 7 – General discussion

7.1 Introduction

Improvement of characteristics associated with animal welfare is of great interest to the livestock industry, due to the increased awareness of welfare aspects in our current production systems among consumers and producers. The research described in this thesis aimed to investigate whether characteristics associated with animal welfare were genetically and genomically determined by using quantitative and molecular genetic approaches and to develop strategies indicating how these traits could be used in breeding programmes. Various approaches for selection for characteristics associated with animal welfare have been discussed in Chapter 2 to Chapter 6. This final chapter provides an overview of the outcomes and discusses the implications and perspectives for future research. In the last section a final conclusion is presented regarding the integration of quantitative and molecular genetic approaches to improve characteristics associated with animal welfare.

7.2 Quantitative genetics

Firstly, piglet survival traits were analysed based on quantitative methods using phenotypic performance and all pedigree information by including the additive genetic relationship matrix in the model to predict genetic and environmental parameters of the traits of interest. For this study, two unique data sets were available: (i) a data set which contained detailed information on a range of reproduction traits, including individual piglet survival and birth weight as well as traits at litter level from five different sire and dam lines and (ii) a data set which contained extensive information on reproduction traits and production traits from a sire line and a dam line that originated from one base population but had been selected based on different breeding goals for more than 25 years.

In the first data set, genetic parameters for piglet survival and weight traits at birth were estimated to gain insight into the efficiency of selection for survival as individual trait of the piglet compared to survival as trait of the sow. Most studies

analysed survival at litter level, and although genetic parameters for survival and weight characteristics at piglet level have been estimated in the past (e.g. GRANDINSON *et al.* 2005; RYDHMER *et al.* 2008), few studies compared estimates at litter level and at individual piglet level. For a binomial trait – such as individual piglet survival at birth – threshold models have been shown to be more appropriate (SORENSEN *et al.* 1995) and this approach was used here using a Bayesian methodology. In addition, studies have shown that birth weight is phenotypically related to piglet survival and may indirectly improve the latter (ROEHE 1999; ROEHE and KALM 2000). Data on individual piglet birth weight and survival allowed the examination of whether additional selection for piglet birth weight could contribute to improvement of piglet survival at birth over and above direct selection for survival.

Selection for piglet survival at birth is expected to be successful, because all lines in Chapter 2 showed considerable variation for this trait and relatively high heritabilities, in particular in lines with low average birth weight. Total heritabilities for survival at individual piglet level were lower, with higher maternal heritabilities than direct heritabilities, similar to what has been reported in the literature (KNOL *et al.* 2002a; SU *et al.* 2008; IBAÑEZ-ESCRICHE *et al.* 2009b; ROEHE *et al.* 2009). Maternal heritabilities of individual birth weight were mostly at moderate magnitude, slightly higher than those reported in literature (ROEHE 1999; KNOL *et al.* 2002a; SU *et al.* 2008) and thus of great interest for selection. For most lines, the correlations between the traits indicated that selection for either individual birth weight or average birth weight would indirectly increase survival at birth. The estimated genetic parameters of weight traits are important to maximise overall genetic improvement in piglet survival and growth and thus to improve the sustainability of pig production.

For pig producers the genetic associations of piglet survival with economically important traits, such as reproduction traits, are of great importance. The gross margin per pig in the UK in 2009 was £ 28.12 (BPEX 2010) and lower litter sizes at

weaning constitute a major loss for pig producers. Studies have shown contradictory results for correlations between survival and litter size traits (SERENIUS *et al.* 2004b; ROSENDO *et al.* 2007b; SU *et al.* 2007). In the present study, unfavourable correlations of survival with the number of piglets born in total were found in most lines, which limit the extent to which both can be increased simultaneously. The unfavourable genetic correlations were especially visible in the dam lines, which may indicate an increase in antagonism between litter size and piglet survival when selection is on the former trait. This emphasises the importance of selection for piglet survival in these lines. However, stabilising selection for correlated traits like birth weight can aid to the improvement of survival. KNAP (2008) reported 10-year genetic trends for the number of piglets born in total, perinatal survival (the complement of stillbirth) and pre-weaning survival in four pig lines from the breeding organisation PIC and showed that selection for traits with a low heritability can be successful. Despite antagonistic correlations between the traits, a positive genetic trend could be achieved in all three traits simultaneously. The Danish pig breeding organisation Danavl reported that selection for litter size at birth led to increased mortality, and subsequently adjusted their breeding goals for dam lines to include selection for litter size at day 5 (reviewed by KNAP 2008). On a similar note, after analysis of piglet survival and its correlations with weight traits and litter size traits in their lines (e.g. KNOL *et al.* 2002b), the Dutch pig breeding organisation Topigs adjusted their breeding goals for dam lines (reviewed by KNAP 2008). All three pig breeding organisations selected for survival, or its complement mortality, at litter level, which may have been due to the increased costs associated with individual piglet measurements but also the inherent difficulty of estimating maternal genetic effects due to confounding with direct genetic effects.

In contrast, genetic correlations between survival and the number of piglets born alive showed favourable estimates and selection for this trait will lead to improvement in survival at birth. In conclusion, heritabilities for survival at birth and reproduction traits were low, but genetic variation was substantial and extensive

pedigree information can be used to improve the accuracy of breeding values, so that genetic improvement is expected to be efficient.

Production traits are of great economic importance for pig producers. Traits such as growth rate, lean meat percentage or backfat thickness are selected for in many breeding programmes, but few studies estimated genetic correlations between these traits and piglet survival. In Chapter 3, genetic correlations between survival and the production traits average daily gain, backfat thickness and muscle depth were estimated to be of low to moderate magnitude. However, an unfavourable genetic correlation between backfat thickness and piglet survival at birth was estimated in the sire line, which may have been a result of the selection pressure on production traits in this line. This correlation suggests that a reduced emphasis of selection for backfat thickness combined with stabilising selection for piglet survival may be beneficial. Overall, some undesirable correlations between survival at birth and reproduction traits or production traits were estimated, but in general these correlations were low so that simultaneous improvement of all traits can be achieved.

Little research has been done to compare sire and dam lines and none investigated how selection changed parameters over years in these lines. Chapter 2 showed that differences between breeds and lines can be considerable and no single strategy is optimal for all breeds. Instead, different strategies for different breeds and lines should be considered. Based on these results it can be concluded that breeding goal differences, with emphasis on reproduction traits in dam lines and on production traits in sire lines, have resulted in different genetic parameters for piglet survival between lines, which agrees with the results of KNOL *et al.* (2002a). The variation in heritabilities found among lines in Chapter 2 indicated that the strategy of selection for an optimal birth weight with lowest variation within litter should be considered per line individually to maximise overall genetic improvement in piglet survival and growth.

In Chapter 3 it was shown how more than 25 years of selection with different breeding goals has changed a sire line and a dam line that originated from one breed – the Large White. As a result of selection, average phenotypic differences between these lines were substantial with 1.5 more piglets born in the dam line and 1.7 mm less backfat thickness in the sire line. Selection pressure on litter size in the dam line may have resulted in the higher undesirable correlation between number of piglets born in total and survival at birth in the dam line as compared to the sire line, whereas selection pressure on production traits in the sire line may have led to the highly undesirable correlation between survival and backfat thickness.

By changing the base population through a combined restriction of the depth of the pedigree and performance data to animals born in 2002 or later, it was possible to investigate how genetic parameters and associations between traits changed within line due to the selection emphasis on different traits. Genetic correlations of reproduction traits and production traits showed more desirable genetic associations, though none of them were significant. Estimates indicated that selection pressure on different traits has altered the heritabilities and correlations of the traits within the line.

The estimated genetic parameters of survival, reproduction and production traits are important to optimise breeding programmes and thus to improve the sustainability of pig production with respect to economics and animal welfare. Both Chapter 2 and Chapter 3 showed that genetic correlations differ substantially between breeds, and within breeds between lines, due to genetic selection, though results may be slightly biased since traits in the breeding goal which are included in this analysis may affect estimates of genetic parameters. Despite similar heritabilities for some traits – for example number born in total – selection response may differ greatly among lines due to differences in correlations among traits. Both analyses indicate the importance of individual selection strategies per line. Chapter 2 analysed the influence of selection for birth weight on survival, while Chapter 3 analysed genetic correlations of survival with production traits. Birth weight is phenotypically associated with

production traits, such as postnatal growth and carcass composition at slaughter (POWELL and ABERLE 1980). Studies have shown that genetic correlations of birth weight with production traits can be favourable (e.g. with daily gain or protein deposition), but also unfavourable (e.g. with backfat, drip loss or intramuscular fat) (HERMESCH *et al.* 2000c; KNOL 2001).

Another important aspect to consider is that, although selection is applied on purebreds in highly regulated environments, the ultimate goal is to improve the performance of the crossbred animals under farm conditions. Studies on selection for litter size traits (ENGBLOM *et al.* 2009) and mortality traits (CECCHINATO *et al.* 2010) have shown that selection for improved performance in purebreds has little influence on the improvement of these traits in crossbred animals. One possible solution could be to use information from crossbred animals to evaluate purebred animals for their crossbred performance using a quantitative approach (NEWMAN *et al.* 2010) or genomic selection (IBÁÑEZ-ESCRICHE *et al.* 2009a).

The estimated genetic associations indicate that genetic improvement of piglet survival can be achieved, but an important aspect is the economic consequence of selection for these traits. As indicated in Chapter 2, selection for individual piglet traits is a viable approach, provided the benefits of higher heritabilities outweigh the added costs of weighing each piglet. Differences in economic values for traits between countries and production systems exist and different production systems may require specifically adapted breeding lines (HANENBERG *et al.* 2010). For the successful implementation of selection for traits associated with animal welfare, studies into the economic aspects of selection for welfare related characteristics are of great benefit. Improvement for traits associated with welfare characteristics can be difficult, due to low to moderate economic values. To improve these traits, producers may have to re-evaluate the economic value of these traits, as for example suggested for selection for disease resistance in pigs (HENRYON *et al.* 2010). Livestock production, including pigs, contributes to greenhouse gas emissions; reduction of emissions without compromising animal welfare or productivity is an important topic

of research. TOMA *et al.* (2008) found that a management strategy in which neonatal survival is increased through improved sow diets, can have a positive impact on the net trade in pig meat and a lower impact on the environment (water and air pollution). Improvement of litter size due to changes in the composition of the diet of the sow reduces the number of sows needed to produce a given quantity of pig meat, which may reduce the environmental footprint of the current pig production systems (reviewed by ASHWORTH *et al.* 2009).

7.3 Molecular genetics

The first chapters of this study focused on using quantitative approaches for the improvement of characteristics associated with animal welfare. Subsequent chapters used a molecular genetic approach to investigate aggressive behaviour in pigs. The approach was validated using behavioural, physiological and weight traits in a well-documented mouse SNP data set. Analysis of this mouse SNP data set showed that genomic selection can provide an increase in predictive ability and accuracy over traditional polygenic selection. Genomic selection showed a high predictive ability in comparison to traditional polygenic selection. It was especially advantageous for traits with lower heritabilities, but also dependant on other factors such as the underlying QTL distribution. In particular in situations where little family information was available, the performance of polygenic selection was low and genomic selection increased the performance considerably. Adding a polygenic effect to the genomic effect did not necessarily improve the predictive ability.

Reducing the number of SNPs did not significantly change the predictive ability for most traits, particularly when selection was within family. Results indicated that an increased density of SNP panels does not always result in equivalent higher efficiency of genomic selection. Depending on the trait, fewer SNPs would be sufficient to ensure consistently high efficiency of genomic selection. Increasing numbers of SNPs available for genomic selection will increase the time and costs required for genomic analyses, and approaches that require fewer computing

resources but perform at an equally high level, such as the mixture approach used here, will become more valuable.

The results from Chapter 4 indicated that there might be an optimum proportion of SNPs to be used for genomic selection, which combines low costs with high efficiency. Depending on a trait's genetic and genomic characteristics such as heritability and QTL structure, a (pre-selected) subset of SNPs may be sufficient and a low-density SNP panel could then be developed for the specific breeding line. An important question that should be addressed is the choice of subset of SNPs for this panel. Genome-wide association studies have shown that a two-stage design with pre-selection of SNPs between steps can reduce costs greatly without reducing the power of the study (e.g. SATAGOPAN and ELSTON 2003; LI 2008) and recently a range of statistical methodologies has been developed that attempt to answer the question for the optimum subset – such as LASSO (USAI *et al.* 2009), elastic net (HARRIS and JOHNSON 2010) and stochastic search variable selection (VERBYLA *et al.* 2009). Different traits will be associated with different subsets of SNPs, and a multi-trait approach should be used to determine the optimum subset of SNPs that allows a sufficiently high coverage of SNPs associated with all traits in the breeding goal.

For the analysis of aggressive behaviour in pigs in chapters 5 and 6, the recently developed PorcineSNP60 panel (62,163 SNPs) from Illumina was used to genotype purebred Yorkshire pigs. Proportions of markers per minor allele frequency were distributed evenly across all possible frequencies, from 0% to 50%. The average LD between adjacent SNPs was 0.36 ± 0.37 with an average distance of 52,659 base pairs. An LD of 0.30 is generally considered to be the minimum level for genomic selection (e.g. SARGOLZAEI *et al.* 2008; BANOS and COFFEY 2010) and was found for markers that were located within 75,000 to 100,000 base pairs of each other or less, with 22,434 SNP pairs having an LD >0.99. Considerable differences were found among the chromosomes for SNP density and linkage disequilibrium. This data set is a valuable source of information on the genomic structure of a pig population. The

extent of LD in this population was substantial, though lower than in the mouse data set used in Chapter 4 (VALDAR *et al.* 2006a), and quickly decayed with increase in distance. Similar patterns of high LD at short range with rapid decay at larger distances were found in cattle (BANOS and COFFEY 2010), sheep (MCRAE *et al.* 2002), horse (WADE *et al.* 2009), rainbow trout (REXROAD and VALLEJO 2009), and dogs (SUTTER *et al.* 2004), with the fastest decrease in LD found in humans (REICH *et al.* 2001). Compared to other species, meaningful LD in pigs was found along a considerable range and decay was less pronounced. This pattern of LD in a population depends on the history of its effective population size. A smaller effective population size means few generations are necessary to arrive at common ancestor alleles, and thus fewer opportunities exist for recombination. As a consequence, LD extends across a larger distance (e.g. SUTTER *et al.* 2004; DE ROOS *et al.* 2008; REXROAD and VALLEJO 2009). This explains the rapid decrease of LD in humans – where the effective population size is large – compared to domesticated species, where domestication and subsequent breed formation caused a decrease in effective population size. Some LD exists at long range within breeds (but not across breeds), but only at short distances does it increase rapidly to high levels of LD (reviewed by GODDARD and HAYES 2009). The comprehensive coverage of the genome as well as the extent of linkage disequilibrium suggest that genome-wide association studies as well as genomic selection are expected to be successful to improve traits, including difficult or costly to measure characteristics such as aggressive behaviour.

The approach used for the analysis of genomic selection for behavioural traits, physiological traits and weight traits in mice was subsequently extended to nine behavioural traits in pigs using this SNP data set. Analysis of the data revealed no QTL with significant effect on nine traits related to aggressive behaviour: three lesion score traits at mixing, three lesion score traits at three weeks post mixing and three behavioural traits (reciprocal aggression, delivery of non-reciprocal aggression and receipt of non-reciprocal aggression). The strategy used in this study investigated the linkage between SNPs and trait variation using a genome-wide association study, but another strategy could be to systematically screen genes known to be involved in

the regulation of the trait, referred to as candidate genes (MORMEDE 2005). TERENINA *et al.* (2010) used a candidate gene approach to genotype pigs that were phenotyped for aggressive behaviour. They focussed on those components of the brain serotonergic system which are involved in the regulation of the dopaminergic and the serotonergic system and the regulation of vasopressin, and found several new polymorphisms.

Studies of a range of traits related to aggressive behaviours in livestock species found less than four significant QTL and up five suggestive QTL in chicken (BUITENHUIS *et al.* 2003a; BUITENHUIS *et al.* 2003b; JENSEN *et al.* 2005; WYSOCKI *et al.* 2010) and pigs (CHEN *et al.* 2009). Research of aggressive behaviour in Drosophila found five candidate genes whose transcripts are involved in a broad range of functions. These range from central nervous system development, metabolism, DNA damage recognition and repair, RNA splicing or mitochondrial transport to various binding processes including calcium ion binding, zinc binding and protein binding (EDWARDS and MACKAY 2009). A study of two inbred mice lines, selected for extreme aggressiveness or extreme non-aggressiveness, found two significant QTL. For both QTL, the possible candidate genes were involved in neurotransmission in the brain (BRODKIN *et al.* 2002).

All traits showed low to moderate heritabilities, and inclusion of genomic information improved the predictive ability for these traits. For models using genomic information, the SNPs captured 55% to 76% of the total genetic variation, similar to the percentage of total genetic variation captured by the genomic models using the mouse data set. When including both a genomic and a polygenic effect, the SNPs captured 31% to 59% of the total genetic variation, while the polygenic effect captured the remaining genetic variance. This is slightly lower than the proportions of the total genetic variance captured by SNPs in the mouse data set, whereby the weight traits in general higher proportions (58 to 73%) captured than the behavioural traits (49 to 61%) or physiological traits (46 to 56%). Simulation studies have shown that, for a trait with a heritability of 0.50, the polygenic effect explained about half of

the total genetic variance, but this increased for a trait with a heritability of 0.10 to 56% to 82% of the total genetic variance (CALUS and VEERKAMP 2007) and that with increasing marker densities the estimated variance of the polygenic component increases (SOLBERG *et al.* 2009b). The number of QTL simulated in these studies may have affected the importance of the polygenic effect, whereby the inclusion of a polygenic effect may be more beneficial when more QTL are present.

Genomic selection showed a consistent performance across different mixture percentages, which indicated that lower numbers of SNPs were still sufficient to predict observations. When no full-sib information was available, accuracy using genomic selection increased considerably for most traits, except lesion score at mixing in the caudal area and reciprocal aggression. Simulation studies have shown that the number of QTL affecting a trait influences the performance of genomic selection, though the influence differed depending on the method that was used (e.g. HABIER *et al.* 2009; KIZILKAYA *et al.* 2009; ZHONG *et al.* 2009). The study by KIZILKAYA *et al.* (2009), for example, found that an increase of the number of QTL explaining a predetermined, constant variance of a trait – which meant less variance attributed to a single QTL – led to a decrease in correlations between simulated and predicted genomic merit. This was also visible in the analysis of the mouse SNP data set, where traits with few large-effect QTL showed a lower PA when selection was between families and to a lesser extent for selection within family. Based on the explained genetic variation, the use of genomic selection on traits related to aggressive behaviour in pigs is expected to be beneficial. Studies have shown that mixing scores are genetically correlated with the three behavioural traits (TURNER *et al.* 2008; TURNER *et al.* 2009). The current study showed that, using genomic selection, lesion scores at mixing resulted in higher predictive abilities than receipt of reciprocal aggression, as well as higher gains in accuracy than reciprocal aggression and delivery of non-reciprocal aggression when comparing genomic selection to polygenic selection. At present, high costs of genotyping may be prohibitive for the use of genomic selection for behavioural traits, but the use of different mixture

distributions indicated that cheaper low-density SNP panels may still be sufficient to achieve a high genetic gain.

Similarly to the analysis of the mouse data, adding a polygenic effect to the model including the genomic effect only slightly improved the predictive ability. However, at present, many breeding companies have phenotypic records spanning several generations, which constitute a valuable source of information. This study was based on a relatively small number of genotyped animals, especially when compared to the large volumes of phenotypic information that breeding companies have collected on many other traits. With this in mind, future genomic selection approaches that can incorporate non-genotyped animals (e.g. CHRISTENSEN and LUND 2010) may show different results regarding the inclusion of a polygenic effect. A worthwhile approach for the improvement of genomic selection is the imputation of missing genotypes or haplotypes (e.g. DING *et al.* 2010; HABIER *et al.* 2010; HICKEY *et al.* 2010; VEREIJKEN *et al.* 2010). This would allow for imputation of genotypes for non-genotyped animals, or alternatively, for animals to be genotyped using a low-density SNP panel, followed by imputation of the unobserved SNPs.

The first part of this thesis focussed on the importance of selection for improved piglet survival as a welfare characteristic. Genomic selection could potentially benefit piglet survival, especially if QTL are available that explain a substantial amount of the genetic variation for piglet survival. At present, genome-wide association studies have detected significant QTL for various reproduction traits in pigs. CASSADY *et al.* (2001) found two significant QTL for the number of stillborn pigs (NSB) and one for the number of fully formed pigs using a single QTL model. Analysis of the same data set with multiple QTL models detected one QTL for the number of fully formed piglets at birth (NN), one for NBA, one for NSB, five for the number of mummified piglets (MUM) and one for birth weight (HOLL *et al.* 2004). In addition, this study identified four paternally expressed QTL for NN and one for MUM, as well as one maternally expressed QTL each for NSB, NN and MUM. A study by CHEN *et al.* (2010) found one suggestive QTL for survival rate of piglets at weaning and three for average weight at weaning. Two significant QTL for ovulation

rate and one for embryo survival have been reported, each explaining up to 4.6% of the phenotypic trait variance (BIDANEL *et al.* 2008), while CASSADY *et al.* (2001) found a significant QTL for ovulation rate on a different chromosome. Overall, more than 50 QTL have been mapped for various reproduction traits (DISTL 2007; HERNANDEZ *et al.* 2009; ONTERU *et al.* 2009). Given the influence of the number and effect-size of QTL on the performance of genomic selection, the number of QTL found for reproductive traits is encouraging and genomic selection for these traits may be beneficial.

Marker assisted selection (MAS) uses markers on the genome to estimate genetic variation between animals at the DNA level. MAS-approaches use markers that (i) code directly for the functional mutation or (ii) are in LD with QTL (GODDARD and HAYES 2009). Gene tests for functional mutations have been commercially available for many species (reviewed by DEKKERS 2004). A well-studied example of a functional mutation for which gene tests are available and used successfully is the ryanodine receptor in the skeletal muscle of pigs, where the mutation leads to meat of inferior quality (LAHUCKY *et al.* 1997; DENBOROUGH 1998). Few markers are known to code directly for a functional mutation because most traits are complex traits, influenced and regulated by many genes. However, targeted study of (regions of) candidate genes may yield markers that can be used for MAS, as for example found in a recent study into candidate regions influencing litter size (BJERRE *et al.* 2010). Similarly, the candidate gene approach used by TERENINA *et al.* (2010) may find polymorphisms associated with aggressive behaviour in pigs which could then be utilised in MAS. A disadvantage of these approaches is that they concentrate on a low number of markers associated with a trait. These markers typically explain only a small proportion of the total genetic variance, thus reducing the predictive ability (GODDARD and HAYES 2009). To avoid the disadvantages of MAS, genomic selection (MEUWISSEN *et al.* 2001) uses markers that cover the whole genome, thereby allowing all QTL to be in LD with at least one marker. However, the uptake of genomic selection in livestock breeding is low, except in dairy cattle breeding (HAYES *et al.* 2009; VANRADEN and SULLIVAN 2010).

MAS has been used successfully in pigs for a few traits, such as meat quality, immune response or disease resistance (reviewed by ROTHSCHILD *et al.* 2007). The results of this study show that genomic selection is likely to be more efficient for traits associated with aggressive behaviour than MAS-approaches, since no large-effect QTL were found. For the traits in the mouse data some QTL were found, and efficiency of MAS versus GS would depend on the amount of genetic variance explained by the QTL.

At present, a multi-trait approach for the software used in this study is in development, and future analysis of the data on aggressive behaviour may be based on a multi-trait model. The combination of welfare related characteristics with other, economically important traits using a multi-trait model may give further insight into the genomic regulation of those traits. This information can be used to achieve a simultaneous improvement of animal welfare and all others breeding goal traits in pig breeding programmes.

7.4 Conclusion

Based on the genetic and genomic analyses, the overall conclusion from this thesis is that selection for traits associated with animal welfare characteristics is expected to be successful. The genetic correlations between piglet survival at birth and birth weight indicate that selection for either individual birth weight or average birth weight would indirectly increase survival. In addition, undesirable genetic correlations between survival at birth and reproduction traits or production traits were found, but generally of low magnitude, so that simultaneous improvement of all traits in the breeding goal can be achieved. This thesis has shown how selection can change genetic parameters over years and how it leads to significant differences between lines originating from one breed. The differences found between breeds and within breeds between lines suggest that breeding organisations should consider selection strategies per line individually to achieve maximum overall improvement.

The use of molecular genetic information in the form of genomic selection shows promising results for selection against aggressive behaviour in pigs. The consistently high performance of genomic selection across models indicates that low-density SNP panels may be sufficient to ensure a high efficiency of genomic selection. This would reduce the high costs associated with genotyping in pig breeding programmes with their short generation interval. An extension of genomic selection methodologies to a multi-trait approach including both genotyped and non-genotyped animals would enable breeding organisations to utilise the large amounts of data collected over the years to improve their breeding programmes. To summarize, this thesis has shown how to optimise quantitative and genomic approaches to improve animal welfare related characteristics efficiently in pig breeding programmes.

References

ALBERT, FW, CARLBORG, O, PLYUSNINA, I, BESNIER, F *et al.*, 2009. Genetic architecture of tameness in a rat model of animal domestication. Genetics 182**:** 541-554.

ALONSO-SPILSBURY, M, RAMÍREZ-NECOECHEA, R, GONZÁLEZ-LOZANO, M, MOTA-ROJAS, D and TRUJILLO-ORTEGA, ME, 2007. Piglet survival in early lactation: a review. Journal of Animal and Veterinary Advances 6**:** 76-86.

AMARAL, AJ, MEGENS, HJ, CROOIJMANS, R, HEUVEN, HCM and GROENEN, MAM, 2008. Linkage disequilibrium decay and haplotype block structure in the pig. Genetics 179**:** 569-579.

ARANGO, J, MISZTAL, I, TSURUTA, S, CULBERTSON, M and HERRING, W, 2005. Threshold-linear estimation of genetic parameters for farrowing mortality, litter size, and test performance of Large White sows. Journal of Animal Science 83**:** 499-506.

AREY, DS, 1999. Time course for the formation and disruption of social organisation in group-housed sows. Applied Animal Behaviour Science 62**:** 199-207.

AREY, DS, and EDWARDS, SA, 1998. Factors influencing aggression between sows after mixing and the consequences for welfare and production. Livestock Production Science 56**:** 61-70.

ASHLEY, PJ, 2007. Fish welfare: current issues in aquaculture. Applied Animal Behaviour Science 104**:** 199-235.

ASHWORTH, CJ, TOMA, LM and HUNTER, MG, 2009. Nutritional effects on oocyte and embryo development in mammals: implications for reproductive efficiency and environmental sustainability. Philosophical Transactions of the Royal Society B-Biological Sciences 364**:** 3351-3361.

BANOS, G, and COFFEY, MP, 2010. Short communication: characterization of the genome-wide linkage disequilibrium in 2 divergent selection lines of dairy cows. Journal of Dairy Science 93**:** 2775-2778.

BAXTER, EM, JARVIS, S, D'EATH, RB, ROSS, DW *et al.*, 2008. Investigating the behavioural and physiological indicators of neonatal survival in pigs. Theriogenology 69**:** 773-783.

BELL, AM, 2005. Behavioural differences between individuals and two populations of stickleback (Gasterosteus aculeatus). Journal of Evolutionary Biology 18**:** 464-473.

BIDANEL, JP, ROSENDO, A, IANNUCCELLI, N, RIQUET, J *et al.*, 2008. Detection of quantitative trait loci for teat number and female reproductive traits in Meishan x Large White F2 pigs. Animal 2**:** 813-820.

BJERRE, D, MARK, T, SORENSEN, P, PROSCHOWSKY, HF *et al.*, 2010. Investigation of candidate regions influencing litter size in Danish Landrace sows. Journal of Animal Science 88**:** 1603-1609.

BOITARD, S, SANSAS, B, SERVIN, B and SANCRISTOBAL, M, 2010. SNP selection using sparse PLS. Communication No. 0160 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

BORGEN, SO, and SKARSTAD, GA, 2007. Norwegian pig farmers' motivations for improving animal welfare. British Food Journal 109**:** 891-905.

BOUQUET, A, CANARIO, L, LIGONESCHE, B and BIDANEL, JP, 2006. Genetic parameters of litter size, piglet preweaning mortality and growth in French Landrace pigs. Communication No. 06-09 in *8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil*.

BPEX, 2008. *Pig Yearbook 2008*. BPEX, Milton Keynes, United Kingdom.

BPEX, 2010. *Pig Yearbook 2010*. BPEX, Milton Keynes, United Kingdom.

BREUER, K, SUTCLIFFE, MEM, MERCER, JT, RANCE, KA *et al.*, 2005. Heritability of clinical tail-biting and its relation to performance traits. Livestock Production Science 93**:** 87-94.

BRODKIN, ES, GOFORTH, SA, KEENE, AH, FOSSELLA, JA and SILVER, LM, 2002. Identification of quantitative trait loci that affect aggressive behavior in mice. The Journal of Neuroscience 22**:** 1165-1170.

BUITENHUIS, AJ, RODENBURG, TB, SIWEK, M, CORNELISSEN, SJB *et al.*, 2003a. Identification of quantitative trait loci for receiving pecks in young and adult laying hens. Poultry Science 82**:** 1661-1667.

BUITENHUIS, AJ, RODENBURG, TB, VAN HIERDEN, YM, SIWEK, M *et al.*, 2003b. Mapping quantitative trait loci affecting feather pecking behavior and stress response in laying hens. Poultry Science 82**:** 1215-1222.

CALUS, MPL, 2009. Genomic breeding value prediction: methods and procedures. Animal 4**:** 157-164.

CALUS, MPL, MEUWISSEN, THE, WINDIG, JJ, KNOL, EF *et al.*, 2009. Effects of the number of markers per haplotype and clustering of haplotypes on the accuracy of QTL mapping and prediction of genomic breeding values. Genetics Selection Evolution (Les Ulis) 41**:** Article No.: 11.

CALUS, MPL, and VEERKAMP, RF, 2007. Accuracy of breeding values when using and ignoring the polygenic effect in genomic breeding value estimation with a marker density of one SNP per cM. Journal of Animal Breeding and Genetics 124**:** 362-368.

CANARIO, L, CANTONI, E, LE BIHAN, E, CARITEZ, JC *et al.*, 2006a. Between-breed variability of stillbirth and its relationship with sow and piglet characteristics. Journal of Animal Science 84**:** 3185-3196.

CANARIO, L, ROY, N, GRUAND, J and BIDANEL, JP, 2006b. Genetic variation of farrowing kinetics traits and their relationships with litter size and perinatal mortality in French Large White sows. Journal of Animal Science 84**:** 1053-1058.

CASSADY, JP, JOHNSON, RK, POMP, D, ROHRER, GA *et al.*, 2001. Identification of quantitative trait loci affecting reproduction in pigs. Journal of Animal Science 79**:** 623-633.

CECCHINATO, A, DE LOS CAMPOS, G, GIANOLA, D, GALLO, L and CARNIER, P, 2010. The relevance of purebred information for predicting genetic merit of survival at birth of crossbred piglets. Journal of Animal Science 88**:** 481-490.

CHEN, CY, GUO, YM, YANG, GC, YANG, ZQ *et al.*, 2009. A genome wide detection of quantitative trait loci on pig maternal infanticide behavior in a large scale

White Duroc x Erhualian resource population. Behavior Genetics 39**:** 213-219.

CHEN, CY, GUO, YM, ZHANG, ZY, REN, J and HUANG, LS, 2010. A whole genome scan to detect quantitative trait loci for gestation length and sow maternal ability related traits in a White Duroc x Erhualian F-2 resource population. Animal 4**:** 861-866.

CHEN, P, BAAS, TJ, MABRY, JW, KOEHLER, KJ and DEKKERS, JCM, 2003. Genetic parameters and trends for litter traits in U.S. Yorkshire, Duroc, Hampshire, and Landrace pigs. Journal of Animal Science 81**:** 46-53.

CHIMONYO, M, DZAMA, K and BHEBHE, E, 2006. Genetic determination of individual birth weight, litter weight and litter size in Mukota pigs. Livestock Science 105**:** 69-77.

CHRISTENSEN, OF, and LUND, MS, 2010. Genomic prediction when some animals are not genotyped. Genetics Selection Evolution 42**:** 8.

CLEVELAND, M, FORNI, S, GARRICK, D and DEEB, N, 2010. Prediction of genomic breeding values in a commercial pig population. Communication No. 0266 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

CRONEY, CC, and MILLMAN, ST, 2007. Board-invited review: the ethical and behavioral bases for farm animal welfare legislation. Journal of Animal Science 85**:** 556-565.

D'EATH, RB, CONINGTON, J, LAWRENCE, AB, OLSSON, IAS and SANDØE, P, 2010. Breeding for behavioural change in farm animals: practical, economic and ethical considerations. Animal Welfare 19**:** 17-27.

D'EATH, RB, ROEHE, R, TURNER, SP, ISON, SH *et al.*, 2009. Genetics of animal temperament: aggressive behaviour at mixing is genetically associated with the response to handling in pigs. Animal 3**:** 1544-1554.

DAETWYLER, HD, WIGGANS, GR, HAYES, BJ, WOOLLIAMS, JA and GODDARD, ME, 2010. Imputation of missing genotypes from sparse to high density using long-range phasing. Communication No. 0539 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

DAMGAARD, LH, RYDHMER, L, LØVENDAHL, P and GRANDINSON, K, 2003. Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. Journal of Animal Science 81**:** 604-610.

DE LOS CAMPOS, G, NAYA, H, GIANOLA, D, CROSSA, J *et al.*, 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182**:** 375-385.

DE ROOS, APW, HAYES, BJ, SPELMAN, RJ and GODDARD, ME, 2008. Linkage Disequilibrium and Persistence of Phase in Holstein-Friesian, Jersey and Angus Cattle. Genetics 179**:** 1503-1512.

DEEB, N, CLEVELAND, MA, FORNI, S, YU, N *et al.*, 2010. Genome-wide association study for scrotal hernia in commercial pig populations. Communication No. 0525 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

DEKKERS, JCM, 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. Journal of Animal Science 82 E-Suppl**:** E313-328.

DENBOROUGH, M, 1998. Malignant hyperthermia. Lancet 352**:** 1131-1136.

DING, XD, ZHANG, Q and SIMIANER, H, 2010. Haplotype reconstruction in half-sib pedigree using a combination of EM algorithm and rule-based methods. Communication No. 0549 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

DISTL, O, 2007. Mechanisms of regulation of litter size in pigs on the genome level. Reproduction in Domestic Animals 42**:** 10-16.

DRUMMOND, H, 2001. A revaluation of the role of food in broodmate aggression. Animal Behaviour 61**:** 517-526.

DU, FX, CLUTTER, AC and LOHUIS, MM, 2007. Characterizing linkage disequilibrium in pig populations. International Journal of Biological Sciences 3**:** 166-178.

DUIJVESTEIJN, N, KNOL, EF, MERKS, JWM, CROOIJMANS, R *et al.*, 2010a. A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. BMC Genetics 11**:** 11.

DUIJVESTEIJN, N, SOARES LOPES, M, VAN HAERINGEN, WA, MERKS, JWM *et al.*, 2010b. Paternal identification in pigs by SNP's. Communication No. 0380 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

EDWARDS, AC, and MACKAY, TFC, 2009. Quantitative trait loci for aggressive behavior in Drosophila Melanogaster. Genetics 182**:** 889-897.

EDWARDS, AC, ROLLMANN, SM, MORGAN, TJ and MACKAY, TFC, 2006. Quantitative genomics of aggressive behavior in Drosophila Melanogaster. Plos Genetics 2**:** 1386-1395.

EDWARDS, SA, 2002. Perinatal mortality in the pig: environmental or physiological solutions? Livestock Production Science 78**:** 3-12.

ENGBLOM, L, LUNDEHEIM, N, SCHNEIDER, MD, DALIN, AM and ANDERSSON, K, 2009. Genetics of crossbred sow longevity. Animal 3**:** 783-790.

FERNÁNDEZ, A, RODRIGÁÑEZ, J, ZUZÚARREGUI, J, RODRÍGUEZ, MC and SILIÓ, L, 2008. Genetic parameters for litter size and weight at different parities in Iberian pigs. Spanish Journal of Agricultural Research 6**:** 98-106.

FERNANDO, RL, and GROSSMAN, M, 1989. Marker assisted selection using best linear unbiased prediction. Genetics Selection Evolution 21**:** 467-477.

FERRAZ, JB, and JOHNSON, RK, 1993. Animal model estimation of genetic parameters and response to selection for litter size and weight, growth, and backfat in closed seedstock populations of large white and Landrace swine. Journal of Animal Science 71**:** 850-858.

GEWEKE, J, 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. pp. 169-193 in *Bayesian Statistics 4* edited by JM BERNARDO, J BERGER, AP DAWID and AFM SMITH. Oxford University Press, Oxford.

GODDARD, ME, and HAYES, BJ, 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 10**:** 381-391.

GORBACH, D, MOTE, B, TOTIR, L, FERNANDO, R and ROTHSCHILD, M, 2010a. Polydactyl inheritance in the pig. Journal of Heredity 101**:** 469-475.

GORBACH, DM, CAI, W, DEKKERS, JCM, YOUNG, JM *et al.*, 2010b. Large-scale SNP association analyses of residual feed intake and its component traits in pigs. Communication No. 0265 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

GRANDINSON, K, 2005. Genetic background of maternal behaviour and its relation to offspring survival. Livestock Production Science 93**:** 43-50.

GRANDINSON, K, LUND, MS, RYDHMER, L and STRANDBERG, E, 2002. Genetic parameters for the piglet mortality traits crushing, stillbirth and total mortality, and their relation to birth weight. Acta Agriculturae Scandinavica Section A Animal Science 52**:** 167-173.

GRANDINSON, K, RYDHMER, L, STRANDBERG, E and SOLANES, FX, 2005. Genetic analysis of body condition in the sow during lactation, and its relation to piglet survival and growth. Animal Science 80**:** 33-40.

GRANDINSON, K, RYDHMER, L, STRANDBERG, E and THODBERG, K, 2003. Genetic analysis of on-farm tests of maternal behaviour in sows. Livestock Production Science 83**:** 141-151.

HABIER, D, FERNANDO, RL and DEKKERS, JCM, 2009. Genomic selection using low-density marker panels. Genetics 182**:** 343-353.

HABIER, D, FERNANDO, RL and GARRICK, DJ, 2010. A combined strategy to infer high-density SNP haplotypes in large pedigrees. Communication No. 0915 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

HANENBERG, EHAT, KNOL, EF and MERKS, JWM, 2001. Estimates of genetic parameters for reproduction traits at different parities in Dutch Landrace pigs. Livestock Production Science 69**:** 179-186.

HANENBERG, EHAT, MATHUR, PK and KNOL, EF, 2010. Marginal economic values for pig production in different countries. Communication No. 0300 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

HARE, B, PLYUSNINA, I, IGNACIO, N, SCHEPINA, O *et al.*, 2005. Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. Current Biology 15**:** 226-230.

HARMEGNIES, N, FARNIR, F, DAVIN, F, BUYS, N *et al.*, 2006. Measuring the extent of linkage disequilibrium in commercial pig populations. Animal Genetics 37**:** 225-231.

HARRIS, BL, and JOHNSON, DL, 2010. SNP selection using elastic net, with application to genomic selection. Communication No. 0282 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

HAYES, B, and GODDARD, ME, 2001. The distribution of the effects of genes affecting quantitative traits in livestock. Genetics Selection Evolution 33**:** 209-229.

HAYES, BJ, BOWMAN, PJ, CHAMBERLAIN, AJ and GODDARD, ME, 2009. Invited review: genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science 92**:** 433-443.

HENRYON, M, SØRENSEN, AC, BERG, P and NIELSEN, B, 2010. Breeding pigs for resistance to disease is difficult even with genomic selection. Communication No. 0854 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

HERMESCH, S, LUXFORD, BG and GRASER, H-U, 2000a. Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs: 1. Description of traits and heritability estimates. Livestock Production Science 65**:** 239-248.

HERMESCH, S, LUXFORD, BG and GRASER, H-U, 2000b. Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs: 2. Genetic relationships between production, carcase and meat quality traits. Livestock Production Science 65**:** 249-259.

HERMESCH, S, LUXFORD, BG and GRASER, H-U, 2000c. Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs: 3. Genetic parameters for reproduction traits and genetic correlations with production, carcase and meat quality traits. Livestock Production Science 65**:** 261-270.

HERNANDEZ, SC, FINLAYSON, HA, ASHWORTH, CJ, HALEY, CS and ARCHIBALD, AL, 2009. Mapping quantitative trait loci for reproduction in pigs. pp. 117-118 in *Control Of Pig Reproduction VIII*, edited by H RODRIQUEZ-MARTINEZ, JL VALLET and AJ ZIECIC. Nottingham University Press, Nottingham.

HICKEY, JM, KINGHORN, BP, CLEVELAND, MA, TIER, B and VAN DER WERF, JHJ, 2010. Recursive long range phasing and long haplotype library imputation: building a global haplotype library for Holstein cattle. Communication No. 0934 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

HILL, WG, and ROBERTSON, A, 1968. Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38**:** 226-231.

HOLL, JW, CASSADY, JP, POMP, D and JOHNSON, RK, 2004. A genome scan for quantitative trait loci and imprinted regions affecting reproduction in pigs. Journal of Animal Science 82**:** 3421-3429.

HUISMAN, AE, CHEREL, P and VAN HAANDEL, B, 2010. Linkage disequilibrium and signatures of selection on chromosome 1 in a commercial sire and dam line. Communication No. 0840 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

IBÁÑEZ-ESCRICHE, N, FERNANDO, RL, TOOSI, A and DEKKERS, JCM, 2009a. Genomic selection of purebreds for crossbred performance. Genetics Selection Evolution 41**:** 10.

IBÁÑEZ-ESCRICHE, N, VARONA, L, CASELLAS, J, QUINTANILLA, R and NOGUERA, JL, 2009b. Bayesian threshold analysis of direct and maternal genetic parameters for piglet mortality at farrowing in Large White, Landrace, and Pietrain populations. Journal of Animal Science 87**:** 80-87.

JANSS, LLG, 2008. *iBay manual version 1.46*. Janss Biostatistics, Leiden, The Netherlands.

JENSEN, P, KEELING, L, SCHUTZ, K, ANDERSSON, L *et al.*, 2005. Feather pecking in chickens is genetically related to behavioural and developmental traits. Physiology and Behavior 86**:** 52-60.

KANIS, E, DE GREEF, KH, HIEMSTRA, A and VAN ARENDONK, JAM, 2005. Breeding for societally important traits in pigs. Journal of Animal Science 83**:** 948-957.

KANIS, E, VAN DEN BELT, H, GROEN, AF, SCHAKEL, J and DE GREEF, KH, 2004. Breeding for improved welfare in pigs: a conceptual framework and its use in practice. Animal Science 78**:** 315-329.

KIZILKAYA, K, FERNANDO, RL and GARRICK, DJ, 2009. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. Journal of Animal Science 88**:** 544-551.

KNAP, PW, 2008. Robustness. pp. 288-301 in *Resource allocation theory applied to farm animal production*, edited by WM RAUW. CAB International, Wallingford.

KNAP, PW, and MERKS, JWM, 1987. A note on the genetics of aggressiveness of primiparous sows towards their piglets. Livestock Production Science 17**:** 161-167.

KNAP, PW, and RAUW, WM, 2008. Selection for high production in pigs. pp. 210-229 in *Resource allocation theory applied to farm animal production*, edited by WM RAUW. CAB International, Wallingford.

KNOL, EF, 2001. Genetic aspects of piglet survival, PhD in *Animal Breeding and Genetics Group*. Wageningen University and Research Centre, Wageningen.

KNOL, EF, DUCRO, BJ, VAN ARENDONK, JAM and VAN DER LENDE, T, 2002a. Direct, maternal and nurse sow genetic effects on farrowing-, pre-weaning- and total piglet survival. Livestock Production Science 73**:** 153-164.

KNOL, EF, LEENHOUWERS, JI and VAN DER LENDE, T, 2002b. Genetic aspects of piglet survival. Livestock Production Science 78**:** 47-55.

LAHUCKY, R, CHRISTIAN, LL, KOVAC, L, STALDER, KJ and BAUEROVA, M, 1997. Meat quality assessed ante- and post mortem by different ryanodine receptor gene status of pigs. Meat Science 47**:** 277-285.

LANGBEIN, J, and PUPPE, B, 2004. Analysing dominance relationships by sociometric methods - a plea for a more standardised and precise approach in farm animals. Applied Animal Behaviour Science 87**:** 293-315.

LAWRENCE, AB, and STOTT, AW, 2009. Profiting from animal welfare: an animal based perspective. *The Oxford Farming Conference 2009, Oxford, United Kingdom*.

LEE, SH, VAN DER WERF, JHJ, HAYES, BJ, GODDARD, ME and VISSCHER, PM, 2008. Predicting unobserved phenotypes for complex traits from whole-genome SNP data. Plos Genetics 4**:** 11.

LEGARRA, A, ROBERT-GRANIE, C, MANFREDI, E and ELSEN, J-M, 2008. Performance of genomic selection in mice. Genetics 180**:** 611-618.

LI, J, 2008. Prioritize and select SNPs for association studies with multi-stage designs. Journal of Computational Biology 15**:** 241-257.

LIINAMO, AE, VAN DEN BERG, L, LEEGWATER, PAJ, SCHILDER, MBH *et al.*, 2007. Genetic variation in aggression-related traits in Golden Retriever dogs. Applied Animal Behaviour Science 104**:** 95-106.

LØVENDAHL, P, DAMGAARD, LH, NIELSEN, BL, THODBERG, K *et al.*, 2005. Aggressive behaviour of sows at mixing and maternal behaviour are heritable and genetically correlated traits. Livestock Production Science 93**:** 73-85.

LUAN, T, WOOLLIAMS, JA, LIEN, S, KENT, M *et al.*, 2009. The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. Genetics 183**:** 1119-1126.

MANOLIO, TA, COLLINS, FS, COX, NJ, GOLDSTEIN, DB *et al.*, 2009. Finding the missing heritability of complex diseases. Nature 461**:** 747-753.

MCRAE, AF, MCEWAN, JC, DODDS, KG, WILSON, T *et al.*, 2002. Linkage disequilibrium in domestic sheep. Genetics 160**:** 1113-1122.

MEUWISSEN, T, 2007. Genomic selection: marker assisted selection on a genome wide scale. Journal of Animal Breeding and Genetics 124**:** 321-322.

MEUWISSEN, THE, 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genetics Selection Evolution (Les Ulis) 41**:** Article No.: 35.

MEUWISSEN, THE, HAYES, BJ and GODDARD, ME, 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157**:** 1819-1829.

MILLIGAN, BN, DEWEY, CE and DE GRAU, AF, 2002. Neonatal-piglet weight variation and its relation to pre-weaning mortality and weight gain on commercial farms. Preventive Veterinary Medicine 56**:** 119-127.

MISZTAL, I, TSURUTA, S, STRABEL, T, AUVRAY, B *et al.*, 2002. BLUPF90 and related programs (BGF90). Communication No. 28-07 in *7th World Congress on Genetics Applied to Livestock Production, Montpellier, France*.

MORGAN-DAVIES, C, WATERHOUSE, A, MILNE, CE and STOTT, AW, 2006. Farmers' opinions on welfare, health and production practices in extensive hill sheep flocks in Great Britain. Livestock Science 104**:** 268-277.

MORMEDE, P, 2005. Molecular genetics of behaviour: research strategies and perspectives for animal production. Livestock Production Science 93**:** 15-21.

MUIR, WM, and CRAIG, JV, 1998. Improving animal well-being through genetic selection. Poultry Science 77**:** 1781-1788.

MULLEN, AM, STAPLETON, PC, CORCORAN, D, HAMILL, RM and WHITE, A, 2006. Understanding meat quality through the application of genomic and proteomic approaches. Meat Science 74**:** 3-16.

MURANI, E, PONSUKSILI, S, D'EATH, RB, TURNER, SP *et al.*, 2010. Association of HPA axis-related genetic variation with stress reactivity and aggressive behaviour in pigs. BMC Genetics 11**:** 74.

NEWMAN, S, WANG, L, ANDERSON, J and CASEY, D, 2010. Utilizing crossbred records to increase accuracy of breeding values in pigs. Communication No. 0266 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

NIELSEN, HM, SONESSON, AK, YAZDI, H and MEUWISSEN, THE, 2009. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. Aquaculture 289**:** 259-264.

OLSEN, HG, HAYES, BJ, KENT, MP, NOME, T *et al.*, 2010. A genome wide association study for QTL affecting direct and maternal effects of stillbirth and dystocia in cattle. Animal Genetics 41**:** 273-280.

ONTERU, SK, FAN, B, GARRICK, DJ, STALDER, KJ and ROTHSCHILD, MF, 2010. Whole-genome association analyses for sow lifetime production, reproduction and structural soundness traits using the PorcineSNP60 BeadChip. Communication No. 0273 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

ONTERU, SK, ROSS, JW and ROTHSCHILD, MF, 2009. The role of gene discovery, QTL analyses and gene expression in reproductive traits in the pig. pp. 87-102 in *Control Of Pig Reproduction VIII*, edited by H RODRIQUEZ-MARTINEZ, JL VALLET and AJ ZIECIC. Nottingham University Press, Nottingham.

PÉREZ-GUISADO, J, LOPEZ-RODRÍGUEZ, R and MUÑOZ-SERRANO, A, 2006. Heritability of dominant-aggressive behaviour in English Cocker Spaniels. Applied Animal Behaviour Science 100**:** 219-227.

PHOCAS, F, BOIVIN, X, SAPA, J, TRILLAT, G *et al.*, 2006. Genetic correlations between temperament and breeding traits in Limousin heifers. Animal Science 82**:** 805-811.

POWELL, SE, and ABERLE, ED, 1980. Effects of birth-weight on growth and carcass composition of swine. Journal of Animal Science 50**:** 860-868.

PRICE, EO, 1984. Behavioral aspects of animal domestication. Quarterly Review of Biology 59**:** 1-32.

PURCELL, S, NEALE, B, TODD-BROWN, K, THOMAS, L *et al.*, 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics 81**:** 559-575.

QUINN, JL, and UETA, M, 2008. Protective nesting associations in birds. Ibis 150**:** 146-167.

QUINTON, VM, WILTON, JW, ROBINSON, JA and MATHUR, PK, 2006. Economic weights for sow productivity traits in nucleus pig populations. Livestock Science 99**:** 69-77.

RAFTERY, AE, and LEWIS, S, 1992. How many iterations in the Gibbs sampler? pp. 763-773 in *Bayesian Statistics 4*, edited by JM BERNARDO, J BERGER, AP DAWID and AFM SMITH. Oxford University Press, Oxford.

RAMOS, AM, CROOIJMANS, R, AFFARA, NA, AMARAL, AJ *et al.*, 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. Plos One 4**:** 13.

RAUW, WM, KANIS, E, NOORDHUIZEN-STASSEN, EN and GROMMERS, FJ, 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. Livestock Production Science 56**:** 15-33.

REICH, DE, CARGILL, M, BOLK, S, IRELAND, J *et al.*, 2001. Linkage disequilibrium in the human genome. Nature 411**:** 199-204.

REXROAD, CE, III, and VALLEJO, RL, 2009. Estimates of linkage disequilibrium and effective population size in rainbow trout. BMC Genetics 10**:** Article No.: 83.

RODENBURG, TB, BIJMA, P, ELLEN, ED, BERGSMA, R *et al.*, 2010. Breeding amiable animals? Improving farm animal welfare by including social effects in breeding programmes. Animal Welfare 19**:** 77-82.

ROEHE, R, 1999. Genetic determination of individual birth weight and its association with sow productivity using Bayesian analyses. Journal of Animal Science 77**:** 330-343.

ROEHE, R, and KALM, E, 2000. Estimation of genetic and environmental risk factors associated with pre-weaning mortality in piglets using generalized linear mixed models. Animal Science 70**:** 227-240.

ROEHE, R, and KENNEDY, BW, 1995. Estimation of genetic parameters for litter size in Canadian Yorkshire and Landrace swine with each parity of farrowing treated as a different trait. Journal of Animal Science 73**:** 2959-2970.

ROEHE, R, SHRESTHA, NP, MEKKAWY, W, BAXTER, EM *et al.*, 2009. Genetic analyses of piglet survival and individual birth weight on first generation data of a selection experiment for piglet survival under outdoor conditions. Livestock Science 121**:** 173-181.

ROSENDO, A, CANARIO, L, DRUET, T, GOGUE, J and BIDANEL, JP, 2007a. Correlated responses of pre- and postweaning growth and backfat thickness to six generations of selection for ovulation rate or prenatal survival in French Large White pigs. Journal of Animal Science 85**:** 3209-3217.

ROSENDO, A, DRUET, T, GOGUE, J, CANARIO, L and BIDANEL, JP, 2007b. Correlated responses for litter traits to six generations of selection for ovulation rate or prenatal survival in French Large White pigs. Journal of Animal Science 85**:** 1615-1624.

ROTHSCHILD, MF, HU, ZL and JIANG, ZH, 2007. Advances in QTL mapping in pigs. International Journal of Biological Sciences 3**:** 192-197.

RYDHMER, L, LUNDEHEIM, N and CANARIO, L, 2008. Genetic correlations between gestation length, piglet survival and early growth. Livestock Science 115**:** 287-293.

SARGOLZAEI, M, SCHENKEL, FS, JANSEN, GB and SCHAEFFER, LR, 2008. Extent of linkage disequilibrium in Holstein cattle in North America. Journal of Dairy Science 91**:** 2106-2117.

SAS, 2002. SAS software, Version 9.1.3 of the SAS System for Windows. pp. SAS Institute Inc., Cary, NC, USA.

SATAGOPAN, JM, and ELSTON, RC, 2003. Optimal two-stage genotyping in population-based association studies. Genetic Epidemiology 25**:** 149-157.

SCHAEFFER, LR, 2006. Strategy for applying genome-wide selection in dairy cattle. Journal of Animal Breeding and Genetics 123**:** 218-223.

SEE, MT, MABRY, JW and BERTRAND, JK, 1993. Restricted maximum likelihood estimation of variance components from field data for number of pigs born alive. Journal of Animal Science 71**:** 2905-2909.

SERENIUS, T, and MUHONEN, P, 2007. Economic values of pork production related traits in Finland. Agricultural and Food Science 16**:** 79-88.

SERENIUS, T, SEVÓN-AIMONEN, M-L, KAUSE, A, MÄNTYSAARI, EA and MAKI-TANILA, A, 2004a. Genetic associations of prolificacy with performance,

carcass, meat quality, and leg conformation traits in the Finnish Landrace and Large White pig populations. Journal of Animal Science 82**:** 2301-2306.

SERENIUS, T, SEVÓN-AIMONEN, M-L, KAUSE, A, MÄNTYSAARI, EA and MÄKI-TANILA, A, 2004b. Selection potential of different prolificacy traits in the Finnish Landrace and Large White populations. Acta Agriculturae Scandinavica Section A Animal Science 54**:** 36-43.

SERENIUS, T, SEVÓN-AIMONEN, M-L and MÄNTYSAARI, EA, 2003. Effect of service sire and validity of repeatability model in litter size and farrowing interval of Finnish Landrace and Large White populations. Livestock Production Science 81**:** 213-222.

SERENIUS, T, and STALDER, KJ, 2004. Genetics of length of productive life and lifetime prolificacy in the Finnish Landrace and Large White pig populations. Journal of Animal Science 82**:** 3111-3117.

SILIÓ, L, FERNÁNDEZ, A, MERCADÉ, A, MARTIN-PALOMINO, P *et al.*, 2010. Measuring inbreeding in a closed pig strain from high-density SNPs genotypes. Communication No. 0480 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

SMALL, MF, 1988. Female primate sexual behavior and conception - are there really sperm to spare. Current Anthropology 29**:** 81-100.

SOLBERG, LC, VALDAR, W, GAUGUIER, D, NUNEZ, G *et al.*, 2006. A protocol for high-throughput phenotyping, suitable for quantitative trait analysis in mice. Mammalian Genome 17**:** 129-146.

SOLBERG, TR, SONESSON, AK, WOOLLIAMS, JA and MEUWISSEN, THE, 2009a. Reducing dimensionality for prediction of genome-wide breeding values. Genetics Selection Evolution 41**:** 8.

SOLBERG, TR, SONESSON, AK, WOOLLIAMS, JA, ODEGARD, J and MEUWISSEN, THE, 2009b. Persistence of accuracy of genome-wide breeding values over generations when including a polygenic effect. Genetics Selection Evolution 41**:** 53.

SORENSEN, DA, ANDERSEN, S, GIANOLA, D and KORSGAARD, I, 1995. Bayesian inference in threshold models using Gibbs sampling. Genetics Selection Evolution 27**:** 229-249.

SOUTHWOOD, OI, and KENNEDY, BW, 1990. Estimation of direct and maternal genetic variance for litter size in Canadian Yorkshire and Landrace swine using an animal model. Journal of Animal Science 68**:** 1841-1847.

SOUZA, CA, RAMAYO, Y, MEGENS, HJ, RODRÍGUEZ, MC *et al.*, 2010. Porcine colonization of the Americas: a 60k SNP story. Communication No. 0510 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

STOTT, AW, MILNE, CE, GODDARD, PJ and WATERHOUSE, A, 2005. Projected effect of alternative management strategies on profit and animal welfare in extensive sheep production systems in Great Britain. Livestock Production Science 97**:** 161-171.

STRANDÉN, I, and VUORI, K, 2006. Relax2: pedigree analysis program. Communication No. 27-30 in *8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil*.

SU, G, GULDBRANDTSEN, B, GREGERSEN, VR and LUND, MS, 2009. Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. Journal of Dairy Science 93**:** 1175-1183.

SU, G, LUND, MS and SORENSEN, D, 2007. Selection for litter size at day five to improve litter size at weaning and piglet survival rate. Journal of Animal Science 85**:** 1385-1392.

SU, G, SORENSEN, D and LUND, MS, 2008. Variance and covariance components for liability of piglet survival during different periods. Animal 2**:** 184-189.

SUTTER, NB, EBERLE, MA, PARKER, HG, PULLAR, BJ *et al.*, 2004. Extensive and breed-specific linkage disequilibrium in Canis Familiaris. Genome Research 14**:** 2388-2396.

TAKEUCHI, Y, HASHIZUME, C, ARATA, S, INOUE-MURAYAMA, M *et al.*, 2009. An approach to canine behavioural genetics employing guide dogs for the blind. Animal Genetics 40**:** 217-224.

TERENINA, E, BAZOVKINA, D, ROUSSEAU, S, SALIN, F *et al.*, 2010. Association between aggressive behavior and candidate gene polymorphisms: study of the brain serotonergic system in pigs. Communication No. 0864 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

TOMA, L, ASHWORTH, CJ and STOTT, AW, 2008. A partial equilibrium model of the linkages between animal welfare, trade and the environment in Scotland. pp. 49-69 in *The 109th EAAE Seminar, Viterbo, Italy*.

TOMA, L, KUPIEC-TEHAN, B, STOTT, AW and REVOREDO-GIHA, C, 2010. Animal welfare, information and consumer behaviour. 1972-1982 in *The 9th European IFSA Symposium, Vienna, Austria*.

TURNER, SP, FARNWORTH, MJ, WHITE, IMS, BROTHERSTONE, S *et al.*, 2006a. The accumulation of skin lesions and their use as a predictor of individual aggressiveness in pigs. Applied Animal Behaviour Science 96**:** 245-259.

TURNER, SP, ROEHE, R, D'EATH, RB, ISON, SH *et al.*, 2009. Genetic validation of postmixing skin injuries in pigs as an indicator of aggressiveness and the relationship with injuries under more stable social conditions. Journal of Animal Science 87**:** 3076-3082.

TURNER, SP, ROEHE, R, MEKKAWY, W, FARNWORTH, MJ *et al.*, 2008. Bayesian analysis of genetic associations of skin lesions and behavioural traits to identify genetic components of individual aggressiveness in pigs. Behavior Genetics 38**:** 67-75.

TURNER, SP, WHITE, IMS, BROTHERSTONE, S, FARNWORTH, MJ *et al.*, 2006b. Heritability of post-mixing aggressiveness in grower-stage pigs and its relationship with production traits. Animal Science 82**:** 615-620.

USAI, MG, GODDARD, ME and HAYES, BJ, 2009. LASSO with cross-validation for genomic selection. Genetics Research 91**:** 427-436.

VAGE, J, and LINGAAS, F, 2008. Single nucleotide polymorphisms (SNPs) in coding regions of canine dopamine- and serotonin-related genes. BMC Genetics 9**:** 8.

VALDAR, W, SOLBERG, LC, GAUGUIER, D, BURNETT, S *et al.*, 2006a. Genome-wide genetic association of complex traits in heterogeneous stock mice. Nature Genetics 38**:** 879-887.

VALDAR, W, SOLBERG, LC, GAUGUIER, D, COOKSON, WO *et al.*, 2006b. Genetic and environmental effects on complex traits in mice. Genetics 174**:** 959-984.

VAN ARENDONK, JAM, VAN ROSMEULEN, C, JANSS, LLG and KNOL, EF, 1996. Estimation of direct and maternal genetic (co) variances for survival within litters of piglets. Livestock Production Science 46**:** 163-171.

VAN DEN BERG, L, VOS-LOOHUIS, M, SCHILDER, MBH, VAN OOST, BA *et al.*, 2008. Evaluation of the serotonergic genes htr1A, htr1B, htr2A, and slc6A4 in aggressive behavior of Golden Retriever dogs. Behavior Genetics 38**:** 55-66.

VANRADEN, P, and SULLIVAN, P, 2010. International genomic evaluation methods for dairy cattle. Genetics Selection Evolution 42**:** 7.

VERBYLA, KL, HAYES, BJ, BOWMAN, PJ and GODDARD, ME, 2009. Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle. Genetics Research 91**:** 307-311.

VEREIJKEN, ALJ, ALBERS, GAA and VISSCHER, J, 2010. Imputation of SNP genotypes in chicken using a reference panel with phased haplotypes. Communication No. 0365 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

VILLUMSEN, TM, JANSS, L and LUND, MS, 2009. The importance of haplotype length and heritability using genomic selection in dairy cattle. Journal of Animal Breeding and Genetics 126**:** 3-13.

VISSCHER, PM, MACGREGOR, S, BENYAMIN, B, ZHU, G *et al.*, 2007. Genome partitioning of genetic variation for height from 11,214 sibling pairs. American Journal of Human Genetics 81**:** 1104-1110.

WADE, CM, GIULOTTO, E, SIGURDSSON, S, ZOLI, M *et al.*, 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 326**:** 865-867.

WILLHAM, RL, 1972. The role of maternal effects in animal breeding: III. biometrical aspects of maternal effects in animals. Journal of Animal Science 35**:** 1288-1293.

WILLHAM, RL, 1980. Problems in estimating maternal effects. Livestock Production Science 7**:** 405-418.

WOLF, J, ZÁKOVÁ, E and GROENEVELD, E, 2008. Within-litter variation of birth weight in hyperprolific Czech Large White sows and its relation to litter size traits, stillborn piglets and losses until weaning. Livestock Science 115**:** 195-205.

WONG, CK, and BERNARDO, R, 2008. Genomewide selection in oil palm: increasing selection gain per unit time and cost with small populations. Theoretical and Applied Genetics 116**:** 815-824.

WYSOCKI, M, STRATZ, P, PREUSS, S and BENNEWITZ, J, 2010. Functional investigation of candidate genes affecting feather pecking in chickens. Communication No. 0254 in *9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany*.

YANG, JA, BENYAMIN, B, MCEVOY, BP, GORDON, S *et al.*, 2010. Common SNPs explain a large proportion of the heritability for human height. Nature Genetics 42**:** 565-U131.

ZHONG, SQ, DEKKERS, JCM, FERNANDO, RL and JANNINK, JL, 2009. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. Genetics 182**:** 355-364.

ZUMBACH, B, MISZTAL, I, TSURUTA, S, HOLL, J *et al.*, 2007. Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. Journal of Animal Science 85**:** 901-908.